



Mathematical Foundations of Signal Processing

Mathematical Foundations of Signal Processing

Benjamín Béjar Haro

Mihailo Kolundžija

Reza Parhizkar

Martin Vetterli

November 24, 2014

Where are we now?

- ➊ Geometrical Tools
 - Hilbert spaces, projections etc.
- ➋ Modeling and Analysis
 - Transforms, etc.
- ➌ Measuring and Processing
 - Sampling and Interpolation
 - *Approximation and Compression*
 - Localization and Uncertainty
 - Compressed Sensing
- ➍ Applications

Approximation and Compression

- 1 Why approximate and compress?
- 2 Approximation of Functions by Polynomials (finite interval)
 - Least square approximation
 - Lagrange interpolation
 - Taylor series expansion
 - Minimax approximation
 - FIR filter design
- 3 Approximation of Functions by Splines
 - B-Splines
 - Shift Invariant Subspaces
 - Expansions in Spline Bases
- 4 Approximation by Series Truncation

Why approximate and compress?

Why dealing with approximation and compression?

In many applications one faces the problem of not being able to exactly represent a measurement (a function, a discrete sequence, a vector) or not being able to exactly transform it for analysis purposes. For example:

- The function may not be known everywhere, but only at a number of specific points. The question is then how to optimally *approximate* the measurement given the available values. E.g., observing a function only on a finite interval, or only at a finite number of points in an interval.
- The function may be observable everywhere, but we may be able to afford to measure it or compute its coefficients at a finite rate for complexity and storage reasons. Then how can we obtain an *approximate* expansion or a *projection* on a subspace (e.g. projection onto bandlimited space, Fourier series truncation).
- Finite precision and storage capability. How can we *quantize* and *compress* the information, *i.e.*, how to efficiently represent it.

Approximation of Functions by Polynomials (finite interval)

Approximation of Functions by Polynomials (finite interval)

Consider a function $x(t)$ describing an physical quantity, that is observed over a finite interval $[a, b]$. A degree- K polynomial

$$p_K(t) = \sum_{k=0}^K \alpha_k t^k,$$

enables to approximate the function

- with a simple parametric representation involving only a few parameters
- based on only a few measurements

Higher degree polynomials are “wigglier” and hence can better approximate “wigglier” functions. The approximation error is given by

$$e_K(t) = x(t) - p_K(t), \quad t \in [a, b].$$

Approximation of Functions by Polynomials (finite interval)

Choosing the approximating polynomial

Various criteria for choosing the approximating polynomials exist. We will study the following cases:

- Minimization of the error in the $\mathcal{L}^2([a, b])$ sense: *Least square approximation*.

$$\text{i.e., minimize } \|e_K\|_2$$

- Matching function values at given points: *Lagrange interpolation*.

$$\text{i.e., set } e_K(t_i) = 0 \text{ at points } t_0, t_1, \dots, t_K$$

- Matching derivatives at a point: *Taylor series expansion*.

$$\text{i.e., set } e_K^{(k)}(t_0) = x^{(k)}(t_0) - p_K^{(k)}(t_0) = 0 \text{ for } k = 0, 1, \dots, K.$$

- Minimization of the error in the $\mathcal{L}^\infty([a, b])$ sense: *Minimax approximation*.

$$\text{i.e., minimize } \|e_K\|_\infty$$

Approximation of Functions by Polynomials (finite interval)

Least square approximation

Here we assume that $x(t) \in \mathcal{L}^2([a, b])$ and we choose the polynomial that minimizes

$$\|e_K\|_2^2 = \int_a^b |x(t) - p_K(t)|^2 dt .$$

Define

$$\mathcal{P}_K([a, b]) = \{\text{polynomials of degree } \leq K \text{ in } [a, b]\} .$$

Observations:

- $\mathcal{L}^2([a, b])$ is a *Hilbert space*
- $\mathcal{P}_K([a, b])$ is a *subspace* of $\mathcal{L}^2([a, b])$!

$$\mathcal{P}_K([a, b]) = \text{span} (\{1, t, t^2, \dots, t^K\})$$

i.e. $p_K(t)$ is of the form $p_K(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \dots + \alpha_K t^K$, for some $\alpha_i \in \mathbb{C}$

- Therefore?

Approximation of Functions by Polynomials (finite interval)

Least square approximation

Use *projection theorem*: The polynomial minimizing the error in the $\mathcal{L}^2([a, b])$ sense is given by the projection of $x(t)$ onto

$$\mathcal{P}_K([a, b]) = \{\text{polynomials of degree } \leq K \text{ in } [a, b]\} \subset \mathcal{L}^2([a, b]).$$

Least square approximation polynomial is given by

$$p_K(t) = \sum_{k=0}^K \langle x, \varphi_k \rangle \varphi_k(t).$$

where $\{\varphi_0(t), \varphi_1(t), \dots, \varphi_K(t)\}$ is an orthonormal basis of $\mathcal{P}_K([a, b])$ (e.g., obtained by Gram-Schmidt orthogonalization).

Approximation of Functions by Polynomials (finite interval)

Orthonormal basis for $\mathcal{P}_K([a, b])$: Legendre polynomials

The Legendre polynomials

$$L_k(t) = \frac{1}{2^k k!} \frac{d^k}{dt^k} (t^2 - 1)^k, \quad k \in \mathbb{N}, t \in [-1, 1],$$

are mutually orthogonal over $[-1, 1]$ and have $\|L_k\| = \sqrt{2/(2k+1)}$.

Hence

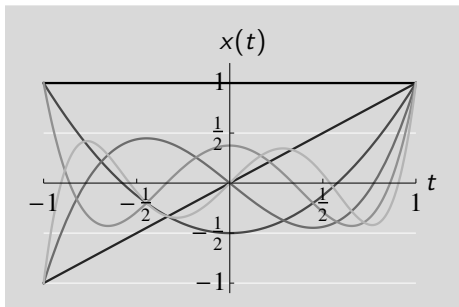
$$\varphi_k(t) = \sqrt{\frac{2k+1}{2}} L_k(t), k = 0, 1, \dots, K$$

forms an orthonormal basis for $\mathcal{P}_K([-1, 1])$!

In fact, shifted and scaled version of L_k forms orthonormal basis for $\mathcal{P}_K([a, b])$.

Approximation of Functions by Polynomials (finite interval)

Least square approximation: Legendre polynomials



The first six Legendre polynomials $\{L_k\}_{k=0}^5$ (from darkest to lightest), which are orthogonal on the interval $[-1, 1]$.

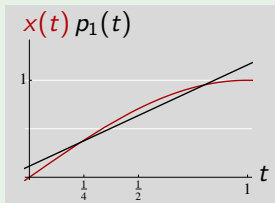
Approximation of Functions by Polynomials (finite interval)

Least square approximation

Approximation with Legendre polynomials

Let's approximate $x(t) = \sin(\pi t/2)$ on $[0, 1]$ with a degree-1 polynomial (a line!). An orthonormal basis of $\mathcal{P}_1([0, 1])$ is $\{1, \sqrt{3}(2t - 1)\}$. So least square approximation polynomial is

$$\begin{aligned} p_1(t) &= \langle \sin(\pi t/2), 1 \rangle 1 + \langle \sin(\pi t/2), \sqrt{3}(2t - 1) \rangle \sqrt{3}(2t - 1) \\ &= \frac{12(4 - \pi)}{\pi^2} t + \frac{8(\pi - 3)}{\pi^2}. \end{aligned}$$



Approximation of Functions by Polynomials (finite interval)

Least square approximation

Higher order approximations

Again let

$$x_1(t) = t \sin 5t.$$

Orthonormal polynomials on $\mathcal{L}^2([0, 1])$ obtained by shifting and scaling the Legendre polynomials:

$$\begin{aligned}\varphi_0(t) &= 1 & \varphi_2(t) &= \sqrt{5}(6t^2 - 6t + 1) \\ \varphi_1(t) &= \sqrt{3}(2t - 1) & \varphi_3(t) &= \sqrt{7}(20t^3 - 30t^2 + 12t - 1)\end{aligned}$$

The best degree- K approximation is

$$p_K(t) = \sum_{k=0}^K \langle x_1, \varphi_k \rangle \varphi_k(t).$$

Approximation of Functions by Polynomials (finite interval)

Least square approximation

Higher order approximations

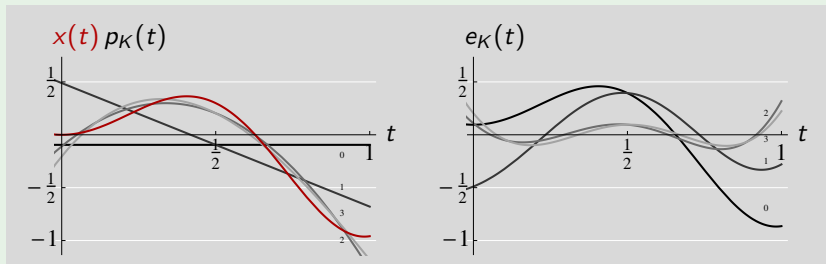
Substituting we get:

$$p_0(t) \approx -0.10$$

$$p_1(t) \approx 0.49 - 1.17t$$

$$p_2(t) \approx -0.11 + 2.42t - 3.59t^2$$

$$p_3(t) \approx -0.20 + 3.56t - 6.43t^2 + 1.90t^3$$



Approximation of Functions by Polynomials (finite interval)

Least square approximation

Same idea can be generalized to other intervals.

The disadvantages of least square approximation: Measurements required for obtaining the approximation parameters, are *inner products* of $x(t)$ with the basis functions. Not always easy to obtain!

Easier measurements: Function values at some points. In other words, *samples* of the function.

Approximation of Functions by Polynomials (finite interval)

Lagrange interpolation

Suppose we observe values of the function $x(t)$ only at $K + 1$ points t_0, \dots, t_K (called *nodes*) and we constrain the approximating polynomial $p_K(t)$ to match these values at these points:

$$p_K(t_i) = x(t_i), \text{ for all } i \in \{0, 1, \dots, K\}.$$

Remark: Remember that a degree- K polynomial has $K + 1$ unknown variables, which justifies the need for $K + 1$ values of $x(t)$.

$$p_K(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \dots + \alpha_K t^K.$$

This leads to a system of $K + 1$ equations with $K + 1$ unknowns

$$\begin{bmatrix} 1 & t_0 & t_0^2 & \cdots & t_0^K \\ 1 & t_1 & t_1^2 & \cdots & t_1^K \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & t_K & t_K^2 & \cdots & t_K^K \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_K \end{bmatrix} = \begin{bmatrix} x(t_0) \\ x(t_1) \\ \vdots \\ x(t_K) \end{bmatrix}$$

A *Vandermonde* system! Invertible if and only if t_i are distinct.

Approximation of Functions by Polynomials (finite interval)

Lagrange interpolation

Unique solution is given by a *Lagrange* polynomial

$$p_K(t) = \sum_{k=0}^K x(t_k) \prod_{i=0, i \neq k}^K \frac{t - t_i}{t_k - t_i}.$$

Theorem (Lagrange interpolation error)

Let $x(t) \in C^{K+1}([a, b])$ ($x(t)$ has $K + 1$ continuous derivatives) and assume the observed points (nodes) t_0, t_1, \dots, t_K are distinct. Then

$$|e_K(t)| \leq \frac{\prod_{k=0}^K |t - t_k|}{(K + 1)!} \max_{\eta \in [a, b]} |x^{(K+1)}(\eta)|.$$

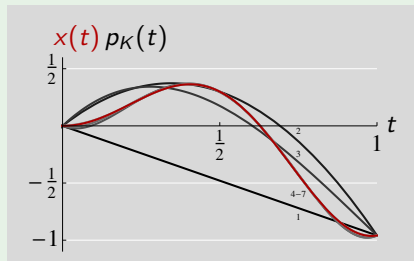
Remark: Notice that, if $x(t)$ is a polynomial of degree K the error is zero, i.e., the interpolant p_K matches everywhere $x(t)$: An expected result since the Lagrange polynomial is unique.

Approximation of Functions by Polynomials (finite interval)

Lagrange interpolation

Lagrange approximation

As before let $x(t) = t \sin 5t$, $t \in [0, 1]$. Choose nodes uniformly as $t_k = \frac{k}{K}$.



Good match over most of the interval, but more error towards the end-points. Matches intuition since bound on error $|e_K(t)|$ is proportional to $\prod_{k=0}^K |t - \frac{k}{K}|$, which tends to be higher near the end-points.

Approximation of Functions by Polynomials (finite interval)

Taylor series expansion

Here we assume $x(t) \in C^K([a, b])$ and find a degree K polynomial that has the same derivatives as function $x(t)$ at some point $t_0 \in [a, b]$.

$$p_K(t) = \sum_{k=0}^K \frac{(t - t_0)^k}{k!} x^{(k)}(t_0).$$

Also called the K -th order *Taylor series approximation* of $x(t)$ around t_0

Theorem (Taylor series expansion error)

Let $x(t) \in C^{K+1}([a, b])$ ($x(t)$ has $K + 1$ continuous derivatives). Then

$$|e_K(t)| \leq \frac{|t - t_0|^{K+1}}{(K + 1)!} \max_{\eta \in [a, b]} |x^{(K+1)}(\eta)|.$$

Observe: Bound on $|e_K(t)|$ proportional to $|t - t_0|^{K+1}$.

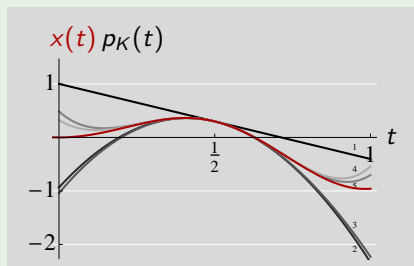
Approximation of Functions by Polynomials (finite interval)

Taylor approximation

Taylor approximation

As before let

$$x(t) = t \sin 5t, t \in [0, 1]$$



Perfect match around $t_0 = 0.5$.

Approximation of Functions by Polynomials (finite interval)

Lagrange vs Taylor

Both Lagrange and Taylor have upper bound on error proportional to

$$\frac{1}{(K+1)!} \max_{\eta \in [a,b]} |x^{(K+1)}(\eta)|.$$

Intuition: Functions that have “higher” magnitude derivatives are “wigglier”, and hence less suitable for approximating with low-degree polynomials.

Recall that functions with higher frequency components need to be sampled at higher rate. Such functions are also “wigglier” and hence less suitable for approximating with low-bandwidth projections.

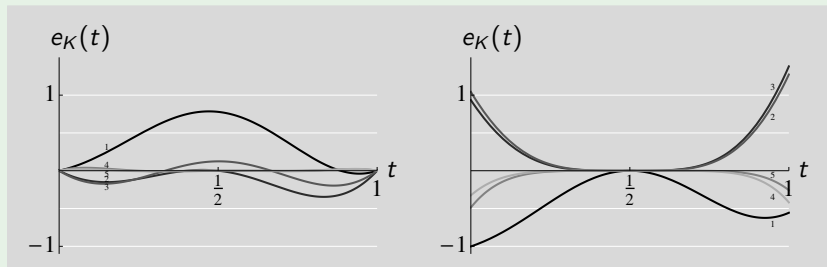
Nevertheless, the nature of Lagrange and Taylor approximations are quite different. Taylor’s expansion is *accurate around t_0* whereas Lagrange’s accuracy *depends on the nodes*.

Approximation of Functions by Polynomials (finite interval)

Lagrange vs Taylor

Who wins?

Approximation for $x(t) = t \sin(5t)$ over $[0, 1]$. Nodes chosen at uniform spacing for Lagrange approximation. For Taylor's $t_0 = 0.5$.



Observe: Lagrange error (figure on left) vanishes at the nodes. Taylor's error (figure on right) is low around 0.5.

Approximation of Functions by Polynomials (finite interval)

Minimax approximation

Here we would like to find the polynomial p_K minimizing

$$\|e_K\|_\infty = \max_{t \in [a,b]} |e_K(t)| = \max_{t \in [a,b]} |x(t) - p_K(t)|.$$

Such a minimization is really not trivial given that WE ARE NOT in a Hilbert space, hence, we cannot exploit Hilbert space tools such as the projection theorem! Finding exact minimax approximation is difficult, however it is easier to approximate!

Approximation of Functions by Polynomials (finite interval)

Minimax approximation

Theorem (Chebyshev equioscillation theorem)

Let $x(t)$ be continuous on $[a, b]$. Let ε_K denote the minimax error when approximating with polynomial of degree at most K . The minimax approximation p_K is unique and determined by the existence of at least $K + 2$ points

$$a \leq s_0 < s_1 < \dots < s_{K+1} \leq b,$$

for which

$$x(s_k) - p_K(s_k) = \sigma(-1)^k \|e_K\|_\infty = \sigma(-1)^k \varepsilon_K,$$

where $\sigma = \pm 1$ is independent of k .

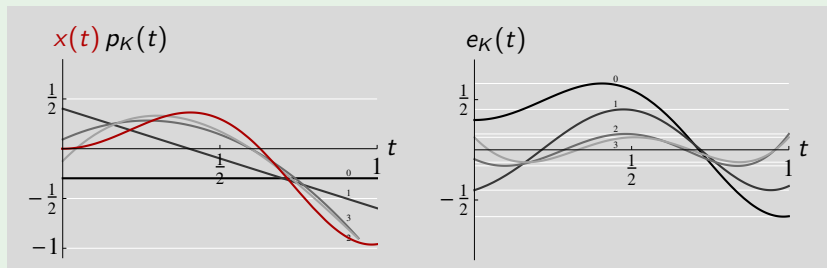
Note: Not a constructive theorem to find the minimax approximation

Approximation of Functions by Polynomials (finite interval)

Minimax approximation: Example

Minimax error

Approximation for $x(t) = t \sin(5t)$ over $[0, 1]$.



Observe: Lower errors for larger K . Equioscillations in error function.

In fact error can be made arbitrarily low for large K .

Approximation of Functions by Polynomials (finite interval)

Theorem (Weierstrass approximation theorem)

Let x be continuous on $[a, b]$ and let $\varepsilon > 0$. Then, there exists a polynomial p for which

$$|e(t)| = |x(t) - p(t)| \leq \varepsilon \quad \text{for every } t \in [a, b].$$

In other words, a *continuous* function can be approximated arbitrarily well on a *finite interval* by a polynomial!

Approximation of Functions by Polynomials (finite interval)

(nearly) Minimax approximation over $[-1, 1]$

If we go back to the Lagrange interpolation, we recall that the error is bounded by

$$|e_K(t)| \leq \frac{\prod_{k=0}^K |t - t_k|}{(K+1)!} \max_{\eta \in [-1, 1]} |x^{(K+1)}(\eta)|$$

Hence maximum error is bounded by

$$\|e_K\|_{\infty} \leq \max_{t \in [-1, 1]} \frac{\prod_{k=0}^K |t - t_k|}{(K+1)!} \max_{\eta \in [-1, 1]} |x^{(K+1)}(\eta)|$$

Idea: choose the nodes t_0, \dots, t_K to minimize right hand side. In other words instead of minimizing the maximum error we are minimizing the maximum bound on the error.

Approximation of Functions by Polynomials (finite interval)

(nearly) Minimax approximation over $[-1, 1]$

The optimal $K + 1$ nodes t_0, \dots, t_K (interpolation points) are given by the roots of a $K + 1$ degree Chebyshev polynomial

$$t_k = \cos \left(\frac{2k + 1}{2(K + 1)} \pi \right), \quad k = 0, 1, \dots, K.$$

Better than choosing interpolation points uniformly. Uniform choice leads to larger errors at the end-points of the interval.

Approximation of Functions by Polynomials (finite interval)

Application to FIR filter design

Given: a desired frequency response H^d assumed to be *real (zero phase) and even*.

Objective: Design an FIR filter with *real coefficients and zero phase*. Assume length $L = 2K + 1$ and *even symmetry*:

$$h_n = \begin{cases} h_{-n}, & \text{for } |n| \leq K; \\ 0, & \text{otherwise.} \end{cases}$$

The frequency response of this filter is

$$\begin{aligned} H(e^{j\omega}) &= \sum_{n=-K}^K h_n e^{-j\omega n} = h_0 + \sum_{n=1}^K h_n (e^{-j\omega n} + e^{j\omega n}) \\ &= h_0 + 2 \sum_{n=1}^K h_n \cos(n\omega). \end{aligned}$$

Note: $H(e^{j\omega})$ is a polynomial of degree n in $t = \cos \omega$. Useful for minimax approximation but not least squares.

Approximation of Functions by Polynomials (finite interval)

FIR filter design: Least-squares approximation

Least-squares criterion

$$\arg \min_{\{h_0, h_1, \dots, h_K\}} \|H^d(e^{j\omega}) - H(e^{j\omega})\|_2^2 = \arg \min_{\{h_0, h_1, \dots, h_K\}} \int_{-\pi}^{\pi} |H^d(e^{j\omega}) - H(e^{j\omega})|^2 d\omega.$$

Equivalently, by Parseval's equality:

$$\arg \min_{\{h_0, h_1, \dots, h_K\}} \|h^d - h\|_2^2 = \arg \min_{\{h_0, h_1, \dots, h_K\}} \sum_{n \in \mathbb{Z}} |h_n^d - h_n|^2.$$

Best solution is to ensure zero contribution from each of the terms

$$n \in \{-K, -K+1, \dots, K\}.$$

Hence *optimal choice* is

$$h_n = h_n^d, \quad \text{for } n = -K, -K+1, \dots, K.$$

Approximation of Functions by Polynomials (finite interval)

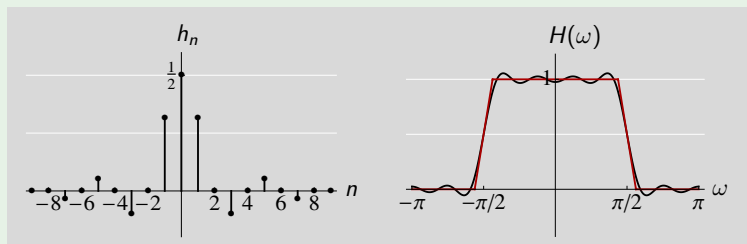
FIR filter design: Least-squares approximation

Ideal lowpass filter

An ideal halfband lowpass filter with unit passband gain:

$$h_n^d = \frac{1}{2} \text{sinc}\left(\frac{1}{2}\pi n\right) \xleftrightarrow{\text{DTFT}} H^d(e^{j\omega}) = \begin{cases} 1, & \text{for } \omega \in [0, \frac{1}{2}\pi]; \\ 0, & \text{for } \omega \in (\frac{1}{2}\pi, \pi]. \end{cases}$$

Choose $K = 7$ (i.e., length 15)



Observe: *Gibbs phenomenon* leads to large absolute error at points of discontinuity.

Approximation of Functions by Polynomials (finite interval)

FIR filter design: Minimax approximation

Minimax criterion in Fourier domain is better justified for FIR filter design.

Theorem (Minimax design criterion)

Let operator $E : \ell^2(\mathbb{Z}) \rightarrow \ell^2(\mathbb{Z})$ denote effect of filtering by difference of h^d and h :

$$Ex = h^d * x - h * x.$$

The operator norm of E is given by (See Solved Exercise 6.3 in book):

$$\|E\| = \max_{\omega \in [-\pi, \pi]} |H^d(e^{j\omega}) - H(e^{j\omega})|.$$

Hence if H^d is real and even, the zero-phase FIR filter h with length $2K + 1$ and even impulse response that minimizes the energy of the difference between filtering by h^d and filtering by h over unit-energy inputs is

$$\arg \min_{\{h_0, h_1, \dots, h_K\}} \max_{\omega \in [0, \pi]} |H^d(e^{j\omega}) - H(e^{j\omega})|.$$

Problem is equivalent to designing a polynomial in $t = \cos \omega$ of degree K . When desired response H^d is continuous *equioscillation theorem* is applicable.

Approximation of Functions by Polynomials (finite interval)

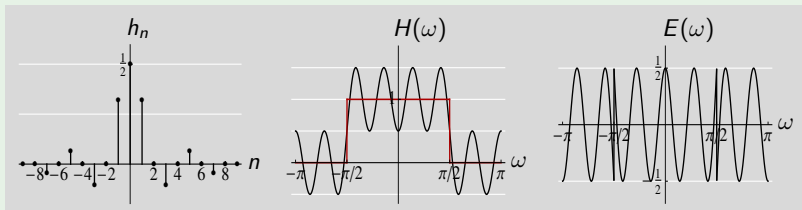
FIR filter design: Minimax approximation

Ideal lowpass filter: Equiripple FIR filter design

An ideal halfband lowpass filter with unit passband gain:

$$h_n^d = \frac{1}{2} \text{sinc}\left(\frac{1}{2}\pi n\right) \xleftrightarrow{\text{DTFT}} H^d(e^{j\omega}) = \begin{cases} 1, & \text{for } \omega \in [0, \frac{1}{2}\pi]; \\ 0, & \text{for } \omega \in (\frac{1}{2}\pi, \pi]. \end{cases}$$

Choose $K = 7$ (i.e., length 15). Solution via Parks-McClellan algorithm (an iterative algorithm that makes use of the equioscillation theorem).



Observe: *Equioscillations* in error function though equioscillation theorem is not applicable since $H^d(e^{j\omega})$ is not continuous.

Approximation of Functions by Polynomials (finite interval)

Recap: Approximation by degree K polynomials

- Minimization of the error in the $\mathcal{L}^2([a, b])$ sense: *Least square approximation*.
 - Solution via projection theorem
 - Can construct orthonormal basis using Legendre polynomials mutually orthogonal over $[-1, 1]$
- Matching function values at $K + 1$ nodes: *Lagrange interpolation*.
 - Inverting a Vandermonde system
- Matching K derivatives at a point t_0 : *Taylor series expansion*.
 - Error is low around t_0 and increases away from t_0
- Taylor and Lagrange: Error bounded by factor proportional to $(K + 1)$ -th derivative of x
 - “Wigglier” functions more difficult to approximate using low degree polynomials
- Minimization of the error in the $\mathcal{L}^\infty([a, b])$ sense: *Minimax approximation*.
 - Chebyshev equioscillation theorem. Exact solution difficult.
 - Use Lagrange method with nodes at roots of Chebyshev polynomial for nearly minimax approximation over $[-1, 1]$.

Approximation of Functions by Polynomials

Advantages and Disadvantages

- + Approximation is smooth
- + Weierstrass theorem: Can approximate continuous functions arbitrarily well over finite intervals
- Cannot approximate discontinuous functions well
- Cannot approximate over infinite interval well
- Approximating continuous functions with high degree polynomials tends to be problematic

Approximation of Functions by Splines

Approximation of Functions by Splines

What are splines?

- A spline function is a function that consists of polynomial pieces joined together with certain smoothness conditions.
- Spline approximation of a function gives a piecewise polynomial approximation which in addition satisfies continuity and smoothness constraints at end-points (called *knots*) of the pieces.
- Advantage over polynomial approximation: Can approximate discontinuous functions well and over the entire real line!

Approximation of Functions by Splines

Definition (Spline, uniform spline, and spline space)

Given $\tau = (\tau_n)_{n \in \mathcal{I}}$, $\tau_n \in \mathbb{R}$, a strictly-increasing sequence, finite or countably infinite.

For $K \in \mathbb{N}$ a function is called a *spline* of degree K with *knots* τ when it is a polynomial of degree at most K on each interval $[\tau_n, \tau_{n+1})$, $n \in \mathcal{I}$, and its derivatives of order $0, 1, \dots, K - 1$ are continuous.

The set of such splines in $\mathcal{L}^2(\mathbb{R})$ is called the *spline space* of degree K with knots τ and is denoted $S_{K,\tau}$.

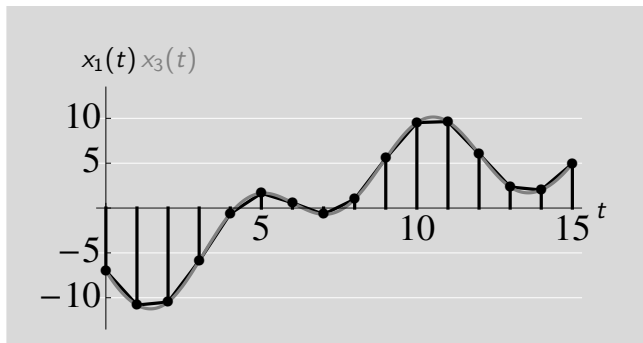
When the knot sequence τ is evenly spaced and doubly infinite, the spline and spline space are called *uniform*.

We will focus on uniform splines.

Qn.: What is $S_{K,\mathbb{Z}}$?

Approximation of Functions by Splines

Splines example



Splines of degree 1 (black line) and 3 (gray line) interpolating the same set of points $\{(t_k, y_k)\}_{k=0}^{15}$.

Splines of degree 1 are piecewise linear and continuous. Splines of degree 3 are piecewise cubic and twice continuously differentiable.

Approximation of Functions by Splines

Degrees of freedom

Consider spline of degree K with $L + 1$ knots.

Have to determine L polynomials of degree K each \Rightarrow have to select $(K + 1)L$ coefficients.

Continuity of function and first $K - 1$ derivatives at $L - 1$ knots $\Rightarrow (L - 1)K$ constraints.

Difference = $(K + 1)L - (L - 1)K = L + K$ degrees of freedom. One per polynomial piece, plus K !

If we specify function values at the knots then we have $L + K - (L + 1) = K - 1$ degrees of freedom remaining.

In practice, for $K = 3$, one specifies derivatives at end-points, or make sure that third derivative is continuous at second and penultimate knots.

Approximation of Functions by Splines

B-Splines

The centered unit-width box function

$$\beta^{(0)}(t) = \begin{cases} 1, & \text{for } t \in [-\frac{1}{2}, \frac{1}{2}); \\ 0, & \text{otherwise,} \end{cases}$$

is called the *elementary B-spline of degree 0*. It is a spline of degree 0 with knots $(-\frac{1}{2}, \frac{1}{2})$. Shifts of $\beta^{(0)}$ are called B-splines of degree 0.

The *elementary B-spline of degree K* is defined by repeated convolution of $\beta^{(0)}$ with itself,

$$\beta^{(K)} = \beta^{(K-1)} * \beta^{(0)}, \quad K = 1, 2, \dots$$

Shifts of $\beta^{(K)}$ are called B-splines of degree K . For example

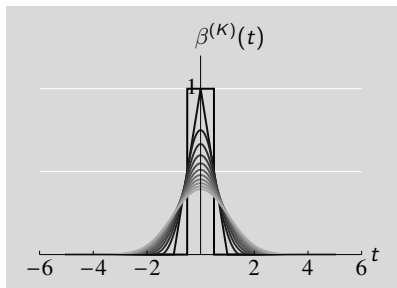
$$\beta^{(1)}(t) = \begin{cases} 1+t, & \text{for } t \in [-1, 0); \\ 1-t, & \text{for } t \in [0, 1); \\ 0, & \text{otherwise,} \end{cases}$$

Fourier transform of $\beta^{(K)}$ is

$$B^{(K)}(\omega) = \text{sinc}^{K+1}\left(\frac{1}{2}\omega\right).$$

Approximation of Functions by Splines

B-Splines

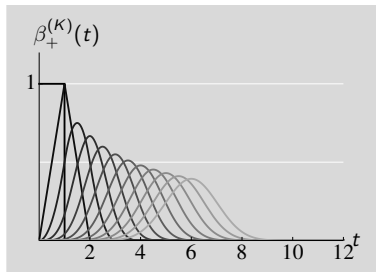


Observations:

- $\beta^{(K)}$ is supported on $[-\frac{K+1}{2}, \frac{K+1}{2})$
- For even K , the function $\beta^{(K)}$ is smooth on intervals of the form $(z - \frac{1}{2}, z + \frac{1}{2})$, $z \in \mathbb{Z}$
- For odd K , the function $\beta^{(K)}$ is smooth on intervals of the form $(z, z + 1)$, $z \in \mathbb{Z}$

Approximation of Functions by Splines

Causal B-Splines



$\beta_+^{(K)}$ are causal versions of $\beta^{(K)}$:

$$\beta_+^{(K)}(t) = \beta^{(K)}\left(t - \frac{1}{2}(K+1)\right), \quad t \in \mathbb{R}.$$

This is called the *causal elementary B-spline of degree K* . Equivalently

$$\beta_+^{(0)}(t) = \begin{cases} 1, & \text{for } t \in [0, 1); \\ 0, & \text{otherwise,} \end{cases}$$

$$\beta_+^{(K)} = \beta_+^{(K-1)} * \beta_+^{(0)}, \quad K = 1, 2, \dots$$

Approximation of Functions by Splines

Shift Invariant Subspaces

Theorem (B-spline bases for uniform spline spaces)

For any $K \in \mathbb{N}$, let $\beta_+^{(K)}$ be the elementary B-spline of degree K and let $\beta_+^{(K)}$ be its causal version. Then, the following statements hold:

- ➊ The causal elementary B-spline $\beta_+^{(K)}$ is a generator of the shift-invariant subspace $S_{K,\mathbb{Z}}$ with respect to shift 1.
- ➋
$$\overline{\text{span}}(\{\beta^{(K)}(t - k)\}_{k \in \mathbb{Z}}) = \begin{cases} S_{K,\mathbb{Z}}, & \text{for odd } K \\ S_{K,\mathbb{Z}+1/2}, & \text{for even } K \end{cases}$$
- ➌ No function with support shorter than that of $\beta_+^{(K)}$ is a generator of $S_{K,\mathbb{Z}}$.

In short: B-splines can be used as building blocks for representing functions in uniform spline spaces! How?

Recall: Sampling followed by ideally matched interpolation leads to orthogonal projection onto a shift-invariant space. Same idea is applicable here.

Approximation of Functions by Splines

Canonical Dual Spline Basis

Let us focus on $K = 1$. We seek *dual basis to* $\{\beta_+^{(1)}(t - k) : k \in \mathbb{Z}\}$. Biorthogonality requires

$$\left\langle \tilde{\beta}_+^{(1)}(t - i), \beta_+^{(1)}(t - k) \right\rangle_t = \delta_{i-k} \quad \text{for every } i, k \in \mathbb{Z}.$$

In addition, for *canonical dual* we need $\tilde{\beta}_+^{(1)} \in S_{1,\mathbb{Z}}$, i.e.,

$$\tilde{\beta}_+^{(1)}(t) = \sum_{k \in \mathbb{Z}} \alpha_k \beta_+^{(1)}(t - k).$$

Exactly same as finding ideally matched interpolation and sampling operators in Chapter 5! In fact, the solved example we saw on finding the sampling prefilter \tilde{g} ideally matched to g is exactly what we need because here $g(t) = \beta_+^{(1)}(t)$.

For any function $x(t)$ the best approximation to x from $S_{1,\mathbb{Z}}$ is given by

$$\hat{x}(t) = \sum_{k \in \mathbb{Z}} \langle x(t), \tilde{\beta}_+^{(1)}(t - k) \rangle_t \beta_+^{(1)}(t - k), \quad t \in \mathbb{R}.$$

Approximation of Functions by Splines

Canonical Dual Spline Basis

Canonical Dual of $\beta^{(1)}(t)$

Since we need $\tilde{\beta}_+^{(1)} \in S_{1,\mathbb{Z}}$ let

$$\tilde{\beta}_+^{(1)}(t) = \sum_{\ell \in \mathbb{Z}} \alpha_\ell \beta_+^{(1)}(t - \ell)$$

Biorthogonality condition becomes

$$\begin{aligned} \delta_k &= \langle \beta_+^{(1)}(t - k), \tilde{\beta}_+^{(1)}(t) \rangle_t \\ &= \sum_{\ell \in \mathbb{Z}} \alpha_\ell \langle \beta_+^{(1)}(t - k), \beta_+^{(1)}(t - \ell) \rangle_t \\ &= \sum_{\ell \in \mathbb{Z}} \alpha_\ell a_{\ell - k} \end{aligned}$$

where a_m denotes autocorrelation sequence

$$a_m = \langle \beta_+^{(1)}(t), \beta_+^{(1)}(t - m) \rangle, \quad m \in \mathbb{Z}.$$

Approximation of Functions by Splines

Canonical Dual of $\beta_+^{(1)}(t)$

To solve for α rewrite as convolution:

$$\delta_k = (\alpha * a_-)_k$$

In z -transform domain, we get

$$\alpha(z)A(z^{-1}) = 1.$$

Substituting $A(z) = (z + 4 + z^{-1})/6$ we get

$$\begin{aligned}\alpha(z) &= \frac{1}{A(z^{-1})} = \frac{6}{z^{-1} + 4 + z} = \frac{6c}{(1 + cz^{-1})(1 + cz)} \\ &= \frac{6c}{1 - c^2} \left(\frac{1}{1 + cz^{-1}} - \frac{cz}{1 + cz} \right),\end{aligned}$$

where $c = 2 - \sqrt{3}$. Inverting we get

$$\alpha_k = \frac{6c}{1 - c^2} (-c)^{|k|}, \quad k \in \mathbb{Z}.$$

Approximation of Functions by Splines

Canonical Dual of $\beta^{(1)}_+(t) \equiv$ Ideally matched prefilter

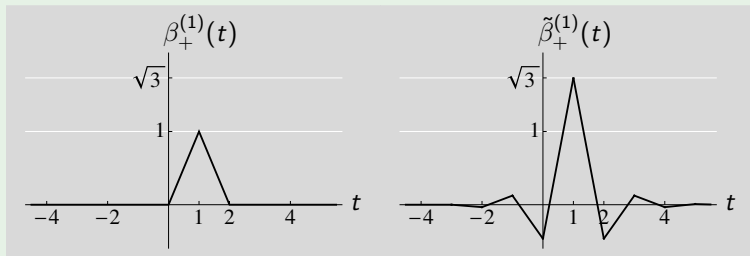


Figure on left: $g(t) = \beta^{(1)}_+(t)$, the generator of $S_{1,\mathbb{Z}}$.

Figure on right: $\tilde{g}(t) = \tilde{\beta}^{(1)}_+(t)$, the canonical dual of $\beta^{(1)}_+(t)$. Can also be interpreted as sampling pre-filter ideally matched to the interpolation post-filter $g(t) = \beta^{(1)}_+(t)$.

Observe: Dual spline has infinite support, but decays quite fast (in fact, exponentially fast).

Approximation of Functions by Splines

Continuous-time operators in discrete-time: Derivatives

To compute the derivative of $x \in S_{K,\mathbb{Z}}$, we can differentiate its series expansion

$$x(t) = \sum_{k \in \mathbb{Z}} \alpha_k \beta_+^{(K)}(t - k), \quad t \in \mathbb{R}, \quad (3)$$

term by term.

Consider first the derivative of $\beta_+^{(0)}(t)$ by using the Dirac's delta

$$\frac{d}{dt} \beta_+^{(0)}(t) = \frac{d}{dt} (u(t) - u(t-1)) = \delta(t) - \delta(t-1),$$

Then, the derivative of $\beta_+^{(K)}(t)$ for $K \in \mathbb{Z}$ can be computed as

$$\begin{aligned} \frac{d}{dt} \beta_+^{(K)}(t) &= \frac{d}{dt} \left(\beta_+^{(K-1)}(t) * \beta_+^{(0)}(t) \right) = \beta_+^{(K-1)}(t) * \frac{d}{dt} \beta_+^{(0)}(t) \\ &= \beta_+^{(K-1)}(t) - \beta_+^{(K-1)}(t-1), \end{aligned}$$

Approximation of Functions by Splines

Continuous-time operators in discrete-time: Derivatives

Then, the derivative of $x \in S_{K,\mathbb{Z}}$ is

$$\begin{aligned}\frac{d}{dt}x(t) &= \frac{d}{dt} \sum_{k \in \mathbb{Z}} \alpha_k \beta_+^{(K)}(t-k) = \sum_{k \in \mathbb{Z}} \alpha_k \frac{d}{dt} \beta_+^{(K)}(t-k) \\&= \sum_{k \in \mathbb{Z}} \alpha_k \beta_+^{(K-1)}(t-k) - \sum_{k \in \mathbb{Z}} \alpha_k \beta_+^{(K-1)}(t-k-1) \\&= \sum_{k \in \mathbb{Z}} \alpha_k \beta_+^{(K-1)}(t-k) - \sum_{\ell \in \mathbb{Z}} \alpha_{\ell-1} \beta_+^{(K-1)}(t-\ell) \\&= \sum_{k \in \mathbb{Z}} (\alpha_k - \alpha_{k-1}) \beta_+^{(K-1)}(t-k) = \sum_{k \in \mathbb{Z}} \alpha'_k \beta_+^{(K-1)}(t-k),\end{aligned}$$

Discrete differentiation

$$\alpha'_k = \alpha_k - \alpha_{k-1}, \quad k \in \mathbb{Z}.$$

Thus, the discrete derivative to the sequence that represents $x(t) \in S_{K,\mathbb{Z}}$ gives the sequence representing $\frac{d}{dt}x(t) \in S_{K-1,\mathbb{Z}}$.

Approximation of Functions by Splines

Continuous-time operators in discrete-time: Integrals

Analog to the derivatives case, integrals of $x \in S_{K,\mathbb{Z}}$ can be computed term by term from its series expansion.

Consider first the integral of $\beta_+^{(K)}(t)$

$$\begin{aligned}\zeta^{(K)}(t) &= \int_{-\infty}^t \beta_+^{(K)}(\tau) d\tau = \sum_{m=0}^{\infty} \int_{t-m-1}^{t-m} \beta_+^{(K)}(\tau) d\tau \\ &= \sum_{m=0}^{\infty} \int_{-\infty}^{\infty} \beta_+^{(K)}(\tau) \beta_+^{(0)}(t-m-\tau) d\tau = \sum_{m=0}^{\infty} \beta_+^{(K+1)}(t-m),\end{aligned}$$

Approximation of Functions by Splines

Continuous-time operators in discrete-time: Integrals

Then, the integral of $x \in S_{K,\mathbb{Z}}$ is

$$\begin{aligned}\int_{-\infty}^t x(\tau) d\tau &= \int_{-\infty}^t \sum_{k \in \mathbb{Z}} \alpha_k \beta_+^{(K)}(\tau - k) d\tau = \sum_{k \in \mathbb{Z}} \alpha_k \int_{-\infty}^t \beta_+^{(K)}(\tau - k) d\tau \\&= \sum_{k \in \mathbb{Z}} \alpha_k \int_{-\infty}^{t-k} \beta_+^{(K)}(s) ds = \sum_{k \in \mathbb{Z}} \alpha_k \sum_{m=0}^{\infty} \beta_+^{(K+1)}(t - k - m) \\&= \sum_{k \in \mathbb{Z}} \alpha_k \sum_{n=k}^{\infty} \beta_+^{(K+1)}(t - n) = \sum_{n \in \mathbb{Z}} \underbrace{\sum_{k=-\infty}^n \alpha_k}_{\alpha_n^{(1)}} \beta_+^{(K+1)}(t - n)\end{aligned}$$

Discrete integration

$$\alpha_k^{(1)} = \sum_{m=-\infty}^k \alpha_m, \quad k \in \mathbb{Z}.$$

Thus, the discrete integral of the sequence that represents $x(t) \in S_{K,\mathbb{Z}}$ gives the sequence representing $\int_{-\infty}^t x(\tau) d\tau \in S_{K+1,\mathbb{Z}}$.

Approximation by Series Truncation

Approximation by Series Truncation

Consider an orthonormal expansion in an infinite dimensional Hilbert space:

$$x = \sum_{k \in \mathbb{Z}} \langle x, \varphi_k \rangle \varphi_k \text{ where } \langle \varphi_k, \varphi_\ell \rangle = \delta_{k-\ell}$$

- Ex. 1: Fourier series expansion of T -periodic functions

$$x(t) = \sum_{k \in \mathbb{Z}} X_k e^{j(2\pi/T)kt}, \quad t \in [-T/2, T/2)$$

where X_k are the Fourier series coefficients

- Here $\varphi_k = \frac{1}{\sqrt{T}} e^{j(2\pi/T)kt}$ and $\{\varphi_k : k \in \mathbb{Z}\}$ forms an orthonormal basis for $\mathcal{L}^2([-T/2, T/2))$
- Ex. 2: Interpolation of bandlimited functions $x \in \text{BL}[-\pi/T, \pi/T]$,

$$x(t) = \sum_{n \in \mathbb{Z}} x(nT) \text{sinc}\left(\frac{\pi}{T}(t - nT)\right), \quad t \in \mathbb{R}.$$

- Here $\varphi_k = \frac{1}{\sqrt{T}} \text{sinc}\left(\frac{\pi}{T}(t - kT)\right)$ and $\{\varphi_k : k \in \mathbb{Z}\}$ forms an orthonormal basis for $\text{BL}[-\pi/T, \pi/T]$
- Ex. 3: Similarly, interpolation of bandlimited sequences

Approximation by Series Truncation

Linear and nonlinear approximations

Given an orthonormal expansion in an infinite dimensional Hilbert space:

$$x = \sum_{k \in \mathbb{Z}} c_k \varphi_k \quad \text{where } c_k = \langle x, \varphi_k \rangle.$$

In practice we cannot store all coefficients. We need to approximate. But how?

Two approaches:

- ➊ Linear approximation: Retain coefficients with an a priori fixed set of indices, e.g., M coefficients around the zeroth coefficient $\{-\frac{(M-1)}{2}, \dots, -1, 0, 1, \dots, \frac{(M-1)}{2}\}$.
 - Choice of indices *does not depend on x*
 - Linear but *not optimal* in error
- ➋ Nonlinear approximation: Retain only M largest coefficients in absolute value
 - Choice of indices *depends on x* (hence is non-linear)
 - *Optimal* in error (Next slide) but: (i) Not linear and (ii) Need to store the indices of the coefficients

Approximation by Series Truncation

Error and optimality of nonlinear approximation

Approximation is given by

$$\hat{x} = \sum_{m \in \mathcal{I}_M} c_m \varphi_m$$

where \mathcal{I}_M denotes the set of coefficients that are kept. Hence squared error is

$$\begin{aligned} \|x - \hat{x}\|^2 &= \left\| \sum_{k \in \mathbb{Z}} c_k \varphi_k - \sum_{m \in \mathcal{I}_M} c_m \varphi_m \right\|^2 \\ &= \left\| \sum_{m \in \mathbb{Z} \setminus \mathcal{I}_M} c_m \varphi_m \right\|^2 = \sum_{m \in \mathbb{Z} \setminus \mathcal{I}_M} |c_m|^2. \end{aligned}$$

Hence error is minimized when \mathcal{I}_M is chosen to be the set of M indices corresponding to the largest coefficients $|c_k|$.

Approximation by Series Truncation

Linear and nonlinear approximations

Fourier series: Linear and nonlinear approximations

Fourier series expansion of box function of period 1, width 1/2 is

$$x(t) = \begin{cases} \sqrt{2}, & \text{for } |t| \leq 1/4; \\ 0, & \text{otherwise,} \end{cases} \quad \xleftrightarrow{\text{FS}} \quad \frac{1}{\sqrt{2}} \operatorname{sinc}(\pi k/2).$$

FS coefficients are symmetric around origin, and all even terms (except at the origin) are zero,

$$X = \sqrt{2} \left[\dots \quad 0 \quad -\frac{1}{3\pi} \quad 0 \quad \frac{1}{\pi} \quad \boxed{\frac{1}{2}} \quad \frac{1}{\pi} \quad 0 \quad -\frac{1}{3\pi} \quad 0 \quad \dots \right].$$

In linear approximation, we retain M FS coefficients centered at the DC coefficient. In non-linear approximation we retain M non-zero coefficients centered at the DC coefficient.

Approximation by Series Truncation

Linear and nonlinear approximations

Fourier series approximation

Using linear approximation, we keep the central $M = 4K - 1$ terms, with squared \mathcal{L}^2 norm of the approximation error

$$\varepsilon_M^2 = \frac{4}{\pi^2} \sum_{|k| \geq K} \frac{1}{(2k+1)^2}.$$

In nonlinear approximation, we skip all the zero terms; this will improve the approximation constant, but not the order, as the squared \mathcal{L}^2 norm of both approximation errors is proportional to $1/M$.

Summary

- Approximation is often necessary when measuring in practice. We studied:
- Approximating functions by polynomials on a finite interval
 - Least square approximation
 - Lagrange interpolation
 - Taylor series approximation
 - Minimax approximation
- Approximating functions by splines
 - Expansion in spline basis is similar to ideally matched sampling and interpolation
 - Continuous operators implemented with discrete-time processing
- Approximation by series truncation: Linear and nonlinear methods

We did not study:

- Compression: Necessary since only finite precision measurements are possible in practice
- Reading: Sections 6.1, 6.2 (excluding 6.2.4), 6.3 (excluding 6.3.3), 6.4.1