Set 15: Orthogonality and Approximation- Part 1

Kyle A. Gallivan

Department of Mathematics

Florida State University

Foundations of Computational Math 1
Fall 2017

Additional References

- J. Dettman, *Mathematical Methods in Physics and Engineering*, McGraw Hill, 1969
- D. Luenberger, *Optimization by Vector Space Methods*, Wiley, 1969

Finite Dimensional Spaces

Let $M = M^*$ be positive definite. $\mathbb{R}^n$ and $\mathbb{C}^n$ are finite dimensional Hilbert spaces with

$$< x, y >= y^* M x, \quad \|x\|_M^2 =< x, x >$$

$$< x, y >=< y, x >^*$$

$$< \alpha x + \beta z, y >= \alpha < x, y > +\beta < z, y >$$

$$\forall x \neq 0, \;\; < x, x > > \; 0, \;\; \text{and} \;\; < x, x >= 0 \leftrightarrow x = 0$$

Vectors x and y are orthogonal when $< x, y >= y^* M x = 0$.

$*$ is transpose or Hermitian transpose.

Finite Dimensional Spaces

A set of orthonormal vectors in $\mathbb{R}^n$ and $\mathbb{C}^n$ satisfy

$$< q_i, q_j >= \delta_{ij}, \quad 1 \leq i, j \leq k \leq n$$

$$Q_k = \begin{pmatrix} q_1 & q_2 & \dots & q_k \end{pmatrix}$$

$$Q_k^* M Q_k = I_k$$

Finite Dimensional Spaces

$\mathcal{R}(Q_k)$ is a k-dimensional subspace with orthonormal basis $\{q_1, q_2, \ldots, q_k\}$

$$v \in \mathcal{R}(Q_k) \leftrightarrow v = \sum_{i=1}^{k} q_i \gamma_i = Q_k c,$$

$$\therefore \gamma_i = < v, q_i > \rightarrow c = Q_k^* M v$$

$$v = \sum_{i=1}^{k} q_i \gamma_i = q_1 < v, q_1 > + q_2 < v, q_2 > + \cdots + q_k < v, q_k >$$

with c unique.

Finite Dimensional Spaces

- angle is preserved from the subspace to $\mathbb{R}^n$ or $\mathbb{C}^n$

$$< v_1, v_2 >=< Q_k c_1, Q_k c_2 >= c_2^* Q_k^* M Q_k c_1 =< c_1, c_2 >= c_1^* c_2$$

- length is preserved

$$\|v\|^2 =< Q_k c, Q_k c >= c^* Q_k^* M Q_k c =< c, c >= \|c\|_2^2 = \sum_{i=1}^{k} |\gamma_i|^2$$

Finite Dimensional Spaces

- $\forall f \in \mathcal{H}, \quad f = Q_n c$, i.e., if $k = n$ then $f \leftrightarrow c$ uniquely.
- $\hat{f} = q_1 < f, q_1 > + q_2 < f, q_2 > + \cdots + q_k < f, q_k > \in \mathcal{R}(Q_k)$

$$\forall v \in \mathcal{R}(Q_k), \quad \|f - \hat{f}\| \leq \|f - v\|$$

$$f - \hat{f} \perp \mathcal{R}(Q_k)$$

- $\forall f \in \mathcal{H}, \quad \|f\|^2 \geq \sum_{i=1}^{k} |\gamma_i|^2$, where $\gamma_i = < f, q_i >, \quad k \leq n$.

An Hilbert Space

Definition 15.1. A Hilbert space is a vector space that:

- has an inner product, denoted $< x, y >$,
- that induces a norm $\|x\|^2 =< x, x >$,
- and is complete under the norm.

Example:

$\mathbb{R}^n$ and $\mathbb{C}^n$ are finite dimensional Hilbert spaces.

An Infinite Dimensional Hilbert Space

If the space is infinite dimensional we must deal with linear combinations that have an infinite number of terms. Convergence?

Definition 15.2. The set $\{b_1, \ldots, b_i, \ldots\}$ where $b_i \in \mathcal{H}$ is an orthogonal sequence if

$$< b_i, b_j > \begin{cases} = 0 & \text{if } i \neq j \\ \neq 0 & \text{if } i = j\ . \end{cases}$$

It is an orthonormal sequence if $< b_i, b_j >= \delta_{ij}$.

An Infinite Dimensional Hilbert Space

Definition 15.3. Given a sequence $\{x_1, \dots, x_i, \dots\}$ with $x_i \in \mathcal{H}$, the infinite series $\sum_{i=1}^{\infty} x_i$ converges to $x \in \mathcal{H}$ if

$$\lim_{n\to\infty} \|s_n - x\| = 0$$

where $s_n = \sum_{i=1}^{n} x_i$ and we assume throughout that the norm is induced by the inner product.

An Infinite Dimensional Hilbert Space

Theorem 15.1. *Given an orthonormal sequence* $\{b_1, \ldots, b_i, \ldots\}$ *with* $b_i \in \mathcal{H}$*, the infinite series* $\sum_{i=1}^{\infty} \gamma_i b_i$ *converges to* $x \in \mathcal{H}$ *if and only if* $\sum_{i=1}^{\infty} |\gamma_i|^2 < \infty$*. Given the convergence, it follows that* $\gamma_i = < x, b_i >$*.*

So, a square summable sequence maps to an element of $\mathcal{H}$.

An Infinite Dimensional Hilbert Space

Lemma 15.2. *(Bessel's Inequality)*

Given an orthonormal sequence $\{b_1, \ldots, b_i, \ldots\}$ *with* $b_i \in \mathcal{H}$*, we have*

$$\forall x \in \mathcal{H} \quad \sum_{i=1}^{\infty} |< x, b_i >|^2 = \sum_{i=1}^{\infty} |\gamma_i|^2 \leq \|x\|^2 < \infty$$

So, $x \in \mathcal{H}$ maps to a square summable sequence.

Infinite Dimensional Hilbert Space

Given an orthonormal sequence $\{b_1, \ldots, b_i, \ldots\}$ with $b_i \in \mathcal{H}$,

$$\sum_{i=1}^{\infty} |\gamma_i|^2 < \infty \Rightarrow x = \sum_{i=1}^{\infty} \gamma_i b_i \Rightarrow \gamma_i =< x,\ b_i >$$

$$x \in \mathcal{H} \Rightarrow \sum_{i=1}^{\infty} |< x,\ b_i >|^2 < \infty \Rightarrow \begin{array}{c} \sum_{i=1}^{\infty} < x,\ b_i > b_i = \hat{x} \in \mathcal{H} \\ \text{and} \\ < x,\ b_i >=< \hat{x},\ b_i > \end{array}$$

Infinite Dimensional Hilbert Space

- When $\mathcal{H}$ is finite dimensional $k = n$ or $k \neq n$ is enough to force $x = \hat{x}$ for all $x \in \mathcal{H}$.
- $\mathcal{H}$ can have subspaces with infinite dimension.
- $x = \hat{x}$ vs $x \neq \hat{x}$ depends on another property of the b_i's

Infinite Dimensional Hilbert Space

Lemma 15.3. *Given a Hilbert space $\mathcal{H}$, let $S \subseteq \mathcal{H}$ (not necessarily a subspace). We have*

- $S^\perp$ *is a closed subspace, i.e., the limit of every convergent sequence on* $S^\perp$ *is in* $S^\perp$.
- $S \subseteq S^{\perp\perp}$
- $S^{\perp\perp}$ *is the smallest closed subspace containing* S*, denoted* $\overline{[S]}$.

An Infinite Dimensional Hilbert Space

Theorem 15.4 (Luenberger, p. 51 and 60)**.** *Let $\mathcal{H}$ be a Hilbert space consider an element $x \in \mathcal{H}$. Given an orthonormal sequence $\{b_1, \ldots, b_i, \ldots\}$ with $b_i \in \mathcal{H}$, the series*

$$\sum_{i=1}^{\infty} < x,\ b_i > b_i$$

converges to an element $\hat{x}$ in the closed subspace, M, generated by the sequence. Further,

$$x - \hat{x} \perp M$$

$$\forall v \in M, \quad \|x - \hat{x}\| \leq \|x - v\|$$

An Infinite Dimensional Hilbert Space

- So, in general, $\hat{x} \neq x$ but it is an optimal approximation.
- This generalizes the finite dimensional classical projection theorem of least squares.
- We will look at specific cases of Theorem 15.4.
- We are interested in subspaces of an infinite dimensional Hilbert space that have finite dimension or fininte codimension.
- We need $\hat{x} = x$ conditions to yield a more desireable infinite expansion that will be truncated to a finite number of terms.

Example

The set of continuous functions on $[0,\ 2\pi]$ with the following inner product

$$< f, g >= \int_0^{2\pi} f(x)g(x)dx$$

is a Hilbert space, $\mathcal{H}$.

Consider the sequence of elements in $\mathcal{H}$

$$b_n(x) = \frac{1}{\sqrt{\pi}} \sin nx, \quad n = 0,\ 1, \ldots$$

Example

$$< b_s, b_r >= \int_0^{2\pi} (\sin sx)(\sin rx) dx = 0, \quad r \neq s$$

$$f = \cos x \in \mathcal{H}$$

$$< f, b_n >= \int_0^{2\pi} (\sin nx)(\cos x) dx = 0$$

The orthonormal sequence $b_0, b_1, \ldots,$ has countably infinite number of orthogonal directions but there is still a direction, given by f, that is orthogonal to **all of them**.

An Infinite Dimensional Hilbert Space

Definition 15.4. An orthonormal sequence, $\{b_i\}$, in a Hilbert space $\mathcal{H}$ is said to be complete if the closed subspace generated by the b_i's is $\mathcal{H}$.

Lemma 15.5. *An orthonormal sequence* $\{b_1, \ldots, b_i, \ldots\}$ *with* $b_i \in \mathcal{H}$ *is complete (or closed) if*

$$\nexists f \in \mathcal{H} \quad | \quad \|f\| = 1 \quad \textit{and} \quad \forall b_i, \;\; < f, b_i >= 0$$

Or equivalently,

$$\forall b_i, \;\; < f, b_i >= 0 \rightarrow f = 0$$

An Infinite Dimensional Hilbert Space

Theorem 15.6. *Given a complete orthonormal sequence* $\{b_1, \ldots, b_i, \ldots\}$ *with* $b_i \in \mathcal{H}$*, we have*

$$\forall f \in \mathcal{H}, \quad \lim_{n \to \infty} \| f - \sum_{i=1}^{n} \gamma_i b_i \|^2 = 0$$

where $\gamma_i =< f, b_i >$.

Additionally, we have the completeness relation or Parseval's equality

$$\| f \|^2 = \sum_{i=1}^{\infty} |\gamma_i|^2$$

Parseval's Equality

Theorem 15.7. *Given a complete orthonormal sequence* $\{b_1, \dots, b_i, \dots\}$ *with* $b_i \in \mathcal{H}$*,*

$$\forall f, g \in \mathcal{H} \quad \textit{let} \quad \gamma_i = < f, b_i > \quad \textit{and} \quad \mu_i = < g, b_i >$$

$$\textit{then} \quad f = \sum_{i=1}^{\infty} \gamma_i b_i, \quad g = \sum_{i=1}^{\infty} \mu_i b_i, \quad \textit{and} \quad < f, g > = \sum_{i=1}^{\infty} \mu_i^* \gamma_i$$

Note. The form in Theorem 15.6 results by taking $f = g$.

Recall, the finite dimensional result

$$< v_1, v_2 > = < Q_k c_1, Q_k c_2 > = c_2^* Q_k^* M Q_k c_1 = < c_1, c_2 > = c_1^* c_2$$

An Infinite Dimensional Hilbert Space

Definition 15.5. Given $\omega(x) : [a, b] \to \mathbb{R}$, a nonnegative integrable function, the space $\mathcal{L}^2_\omega[a, b]$ is the set

$$\left\{ f \mid f : [a, b] \to \mathcal{S}, \quad \int_a^b |f(x)|^2 \omega(x) dx < \infty \right\}$$

$$\mathcal{S} = \mathbb{R} \quad \text{or} \quad \mathcal{S} = \mathbb{C}$$

The inner product and induced norm on the space are

$$(f, g)_\omega = \int_a^b g(x)^* f(x) \omega(x) dx \quad \text{and} \quad \|f\|_\omega^2 = \int_a^b |f(x)|^2 \omega(x) dx$$

An Infinite Dimensional Hilbert Space

- The function $\omega(x)$ is analogous to the positive definite matrix M.
- The integration here is in general Lebesgue.
- Riemann integration can be used when f is taken to be piecewise continuous on $[a, b]$ with, e.g., a finite number of discontinuities.
- Equality is "equal almost everywhere" so the 0 element is an equivalence class of functions that are 0 everywhere but, e.g., a finite number of points for Riemannian integration and more generally a set of measure 0 for Riemannian and Lebesgue integration.
- Convergence in $\|f\|_\omega^2$ is called convergence in the mean.
- Convergence in the mean $\not\Longrightarrow$ pointwise convergence

An Infinite Dimensional Hilbert Space

- When $\mathcal{S} = \mathbb{C}$ care must be taken to be consistent in the inner product and use of the complex conjugate,

$$\forall f, g \in \mathcal{L}^2_\omega[a,b], \quad (f,g)_\omega = (g,f)^*_\omega$$
$$\forall \alpha \in \mathbb{C}, \quad |\alpha|^2 = \alpha^*\alpha = \alpha\alpha^*$$

- The text considers $\mathcal{S} = \mathbb{R}$, i.e., real-valued functions on the real interval $[a,b]$, so we have

$$\forall f, g \in \mathcal{L}^2_\omega[a,b], \quad (f,g)_\omega = (g,f)_\omega = \int_a^b f(x)g(x)\omega(x)dx$$
$$\forall \alpha \in \mathbb{R}, \quad |\alpha|^2 = \alpha^2$$

Representation on the Space

Functions in the space $\mathcal{L}^2_\omega[a,b]$ can be represented via bases with a countably infinite number of elements:

$\forall f \in \mathcal{L}^2_\omega[a,b]$ we have

$$f(x) = \sum_{i=0}^{\infty} \alpha_i \phi_i(x)$$

where

$$\{\phi_0(x), \phi_1(x), \ldots\}$$

is a complete set of orthogonal functions, i.e., a basis for $\mathcal{L}^2_\omega[a,b]$.

Orthogonal Basis and the Representation

Definition 15.6. Let $\{\phi_0(x), \phi_1(x), \ldots\}$ be a complete set of orthogonal functions in $\mathcal{L}^2_\omega[a,b]$. $\forall f \in \mathcal{L}^2_\omega[a,b]$ the series

$$Sf = \sum_{i=0}^{\infty} \alpha_i \phi_i(x) \text{ where } \alpha_i = \frac{(f, \phi_i)_\omega}{(\phi_i, \phi_i)_\omega}$$

is called the generalized Fourier series of f.

In terms of an orthonormal sequence $\{\tilde{\phi}_0(x), \ldots\}$,

$$Sf = \sum_{i=0}^{\infty} \tilde{\alpha}_i \tilde{\phi}_i(x), \quad \tilde{\phi}_i = \frac{\phi_i}{\|\phi_i\|}, \quad \tilde{\alpha} = (f, \tilde{\phi}_i)_\omega = \alpha\|\phi_i\|$$

Orthogonal Polynomials

- choose $[a, b]$ and $\omega(x)$
- Gram-Schmidt process applied to $\{1, x, x^2, \ldots\}$
- low-order recurrence relation from basic properties
- orthogonality and inner product values
- Rodrigues' form – derivative of a polynomial
- some $\phi_i(x)$ are eigenfunctions of Sturm-Liouville differential equation for a specific set of coefficients $\rightarrow$ they form an orthogonal basis, e.g. Jacobi polynomials.

Sturm-Liouville Theory

Definition 15.7. The regular Sturm-Liouville differential equation on $a \le x \le b$ is

$$-(p(x)u'(x))' + q(x)u(x) = \lambda\omega(x)u(x)$$

$$\alpha_0 u'(a) - \alpha_1 u(a) = 0$$

$$\beta_0 u'(b) - \beta_1 u(b) = 0$$

$$p(x) \in \mathcal{C}^1[a,b], \quad w(x), q(x) \in \mathcal{C}^0[a,b]$$

$$a < x < b, \quad p(x) > 0, \;\; \omega(x) > 0, \;\; q(x) \ge 0$$

where at least one of the α pair and at least one of the β pair are nonzero. $(\lambda, u(x))$ pairs are an eigenvalue and its associated eigenfunction.

Sturm-Liouville Theory

Definition 15.8. The singular Sturm-Liouville differential equation on $a \leq x \leq b$ is

$$-(p(x)u'(x))' + q(x)u(x) = \lambda\omega(x)u(x)$$

$$p(x) \in \mathcal{C}^1[a, b], \quad w(x), q(x) \in \mathcal{C}^0[a, b]$$

$$a < x < b, \quad p(x) > 0, \;\; \omega(x) > 0, \;\; q(x) \geq 0$$

with $p(a) = p(b) = 0$ and $p(x)u'(x) \to 0$ as $x \to a$ or as $x \to b$. In other words, u' cannot grow faster than p goes to 0 at the boundary.

Sturm-Liouville Theory

Theorem 15.8. *Given the Sturm-Liouville differential equation on* $a \leq x \leq b$*, there exists a countably infinite set of eigenvalues and associated eigenfunctions* $(\lambda_i, u_i(x))$*,* $i = 1, 2, \ldots$ *such that*

- $0 \leq \lambda_1 < \lambda_2 < \lambda_3 < \cdots$
- $\forall i \neq j, \;\; (u_i, u_j)_\omega = 0$
- $(u_i, u_i)_\omega \neq 0$
- $\{u_1(x), u_2(x), \ldots\}$ *is a complete orthogonal set*
- $u_i(x)$ *has* $i - 1$ *distinct* 0*'s in* $a < x < b$

Legendre Polynomials

- $[a, b] = [-1, 1]$ and $\omega(x) = 1$
- $P_0(x) = 1$, $P_1(x) = x$ and
$$P_{n+1}(x) = \frac{2n+1}{n+1} x P_n(x) - \frac{n}{n+1} P_{n-1}(x)$$
- $P_n(1) = 1$ and $P_n(-x) = (-1)^n P_n(x)$
- Rodrigues' form
$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n]$$
- orthogonality: $(P_n, P_m) = 0$ for $m \neq n$ and
$$(P_n, P_n) = \frac{2}{2n+1}$$

Legendre and the Singular Sturm-Liouville Equation

$$-1 \leq x \leq 1, \quad p(x) = 1 - x^2, \quad q(x) = 0, \quad \omega(x) = 1$$

$$p(1) = p(-1) = 0, \quad \lambda = n(n+1)$$

$$\left[(1 - x^2)y'\right]' + \lambda y = 0$$

$$(1 - x^2)y'' - 2xy' + n(n+1)y = 0$$

Easy to verify the S-L problem is satisfied for

$$P_0(x) = 1, \;\; P_1(x) = x, \;\; P_2(x) = \frac{1}{2}(3x^2 - 1), \;\; \ldots$$

Chebyshev Polynomials

- $[a, b] = [-1, 1]$ and $\omega(x) = 1/\sqrt{1-x^2}$
- $T_n(x) = \cos(n \arccos x)$
- $T_0(x) = 1$, $T_1(x) = x$ and

$$T_{n+1} = 2xT_n(x) - T_{n-1}(x)$$

- $T_n(1) = 1$ and $T_n(-x) = (-1)^n T_n(x)$
- Rodrigues' form $n \geq 1$

$$T_n(x) = \frac{\sqrt{1-x^2}}{(-1)^n(2n-1)(2n-3)\cdots 1} \frac{d^n}{dx^n}[(1-x^2)^{n-1/2}]$$

- orthogonality: $(T_n, T_m) = 0$ for $m \neq n$ and

$$(T_0, T_0) = \pi \text{ and } (T_n, T_n) = \frac{\pi}{2}, \; n \geq 1$$

Chebyshev and the Singular Sturm-Liouville Equation

$$-1 \leq x \leq 1, \quad p(x) = \sqrt{1-x^2}, \quad q(x) = 0, \quad \omega(x) = 1/\sqrt{1-x^2}$$

$$p(1) = p(-1) = 0, \quad \lambda = n^2$$

$$\left[\sqrt{1-x^2}\ \ y'\right]' + \frac{n^2}{\sqrt{1-x^2}} y = 0$$

$$(1-x^2)y'' - xy' + n^2 y = 0$$

Easy to verify the S-L problem is satisfied for

$$T_0 = 1, \quad T_1 = x, \quad , T_2 = 2x^2 - 1 \ldots$$

Jacobi Polynomials

The algebraic polynomials that are eigenfunctions of the singular Sturm-Liouville equation are the two-parameter Jacobi polynomials:

$$J^{\alpha\beta}(x), \quad \alpha > -1, \quad \beta > -1$$

$$p(x) = (1-x)^{\alpha+1}(1+x)^{\beta+1}, \;\; q(x) = 0, \;\; \omega(x) = (1-x)^{\alpha}(1+x)^{\beta}$$

$$p(1) = p(-1) = 0, \quad \lambda = n(n+\alpha+\beta+1)$$

$$(p(x)y'(x))' + \lambda\omega(x)y(x) = 0$$

$$\alpha = \beta = 0 \rightarrow \quad \text{Legendre polynomials}$$

$$\alpha = \beta = -\frac{1}{2} \rightarrow \quad \text{Chebyshev polynomials}$$

Laguerre Polynomials

- normalized form
- $[a, b] = [0, \infty)$ and $\omega(x) = e^{-x}$
- $L_0(x) = 1$, $L_1(x) = -x + 1$ and

$$(n+1)L_{n+1} = (2n+1-x)L_n(x) - nL_{n-1}(x)$$

- Rodrigues' form $n \geq 1$

$$L_n(x) = \frac{e^x}{n!}\frac{d^n}{dx^n}[x^n e^{-x}]$$

- orthogonality: $(L_n, L_m) = 0$ for $m \neq n$ and

$$(L_n, L_n) = 1$$

Laguerre and the Singular Sturm-Liouville Equation

$$0 \le x \le \infty, \quad p(x) = xe^{-x}, \quad q(x) = 0, \quad \omega(x) = e^{-x}$$

$$p(0) = p(\infty) = 0, \quad \lambda = n$$

$$\left[xe^{-x} \;\; y'\right]' + ne^{-x}y = 0$$

$$xy'' + (1-x)y' + ny = 0$$

Easy to verify the S-L problem is satisfied for

$$L_0 = 1, \quad L_1 = 1 - x, \quad , L_2 = \frac{1}{2}(x^2 - 4x + 2) \ldots$$

Hermite Polynomials

- $[a, b] = (-\infty, \infty)$ and $\omega(x) = e^{-\sigma x^2}$
- $\sigma = 1$ is physics form and $\sigma = 1/2$ is probability form
- $H_0(x) = 1$, $H_1(x) = 2\sigma x$ and

$$H_{n+1} = 2\sigma x H_n(x) - 2n\sigma H_{n-1}(x)$$

- $H_n(-x) = (-1)^n H(x)$
- Rodrigues' form $n \geq 0$, $\sigma = 1$

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} [e^{-x^2}]$$

- orthogonality: $\sigma = 1$, $(H_n, H_m) = 0$ for $m \neq n$ and $(H_n, H_n) = 2^n n! \sqrt{\pi}$.

Hermite and the Singular Sturm-Liouville Equation

$$\sigma = 1, \;\; -\infty \le x \le \infty, \quad p(x) = e^{-x^2}, \quad q(x) = 0, \quad \omega(x) = e^{-x^2}$$

$$p(-\infty) = p(\infty) = 0, \quad \lambda = 2n$$

$$\left[e^{-x^2} \;\; y'\right]' + 2ne^{-x^2} y = 0$$

$$y'' - 2xy' + 2ny = 0$$

Easy to verify the S-L problem is satisfied for

$$H_0 = 1, \quad H_1 = 2x, \;\; , H_2 = 4x^2 - 2 \ldots$$

Fourier Polynomials

- $[a, b] = (0, 2\pi)$ and $\omega(x) = 1$
- complex-valued $f(x)$ used to define space:

$$\mathcal{L}^2_\omega[a, b] = \left\{ f \mid f : [a, b] \to \mathbb{C}, \quad \int_a^b |f(x)|^2 dx < \infty \right\}$$

- The inner product and induced norm on the space are

$$(f, g)_\omega = \int_a^b g(x)^* f(x) dx \text{ and } \|f\|^2_\omega = \int_a^b |f(x)|^2 dx$$

- $\phi_k(x) = e^{ikx}$ where $i = \sqrt{-1}$ for $k = 0, \pm 1, \pm 2, \ldots$
- orthogonality: $(\phi_n, \phi_m) = 0$ for $m \neq n$ and $(\phi_n, \phi_n) = 2\pi$
- trigonometric polynomials: $e^{ikx} = \cos kx + i \sin kx$
- related to regular Sturm-Liouville ODE

Orthogonal Polynomials

Suppose $\{\phi_0(x), \phi_1(x), \ldots\}$ is a complete set of orthogonal polynomials in $\mathcal{L}^2_\omega[a, b]$.

Theorem 15.9. *The roots, $x_1, \ldots, x_n$ of $\phi_n(x)$ are real, simple and lie in the interval $a < x < b$, i.e., the interior of $[a, b]$.*

This was mentioned with the Sturm-Liouville Theory characterization of complete orthogonal eigenfunctions.

It can be proven directly based on orthogonality.

Orthogonal Polynomials

Theorem 15.10. *Suppose* $\{\phi_0(x), \phi_1(x), \ldots\}$ *is a complete set of orthogonal polynomials in* $\mathcal{L}^2_\omega[a,b]$ *with* $(\phi_i, \phi_i) = 1$. *If* α_k *and* β_k *are the coefficients of of* x^k *and* x^{k-1} *respectively in* $\phi_k(x)$ *then*

$$\phi_{n+1}(x) = (a_n x + b_n)\phi_n(x) - c_n \phi_{n-1}(x)$$

$$a_n = \frac{\alpha_{n+1}}{\alpha_n}, \quad b_n = \frac{\alpha_{n+1}}{\alpha_n}\left(\frac{\beta_{n+1}}{\alpha_{n+1}} - \frac{\beta_n}{\alpha_n}\right), \quad c_n = \frac{\alpha_{n+1}\alpha_{n-1}}{\alpha_n^2}$$

Orthogonal Polynomials

The recurrence can be initialized with

- $\phi_{-1}(x) = 0$ and $\phi_0(x) = \alpha_0$
- or with $\phi_0(x)$ and $\phi_1(x)$ specific polynomials.

In the latter case take note of the normalization of the norm of $\phi_i(x)$, e.g., it is often made 1 but not necessarily.

The textbook has a form of the recurrence that assumes monic $\phi_i(x)$.

Truncated Generalized Fourier Series

Lemma. *Let $f \in \mathcal{L}^2_\omega[a,b]$ and define the generalized Fourier series truncated after $n+1$ terms*

$$f_n(x) = \sum_{i=0}^{n} \alpha_i \phi_i(x).$$

We have convergence in the mean, i.e., in the $\mathcal{L}^2_\omega[a,b]$ sense,

$$\lim_{n\to\infty} \|f - f_n\|_\omega = 0.$$

Further,

$$\|f\|^2_\omega = \sum_{i=0}^{\infty} \alpha_i^2 \|\phi_i(x)\|^2_\omega, \quad \|\hat{f}\|^2_\omega = \sum_{i=0}^{n} \alpha_i^2 \|\phi_i(x)\|^2_\omega$$

$$\|f - f_n\|^2_\omega = \sum_{i=n+1}^{\infty} \alpha_i^2 \|\phi_i(x)\|^2_\omega$$

Optimality

Theorem 15.11. *Let* $\{\phi_0(x), \phi_1(x), \ldots\}$ *be a complete set of orthogonal functions in* $\mathcal{L}^2_\omega[a,b]$*,* $f \in \mathcal{L}^2_\omega[a,b]$ *and* $f_n(x) = \sum_{i=0}^{n} \alpha_i \phi_i(x)$*.*

If $\mathcal{S}_n = span[\phi_0(x), \phi_1(x), \ldots, \phi_n(x)]$ *then*

$$\|f - f_n\|_\omega = \min_{q \in \mathcal{S}_n} \|f - q\|_\omega$$

Optimality

Proof. We have

$$r_n = f - f_n = \sum_{i=n+1}^{\infty} \alpha_i \phi_i(x)$$

$$\therefore \quad r_n \perp \phi_j, \quad 0 \le j \le n \quad \text{and} \quad \forall q \in \mathcal{S}_n \quad r_n \perp q$$

We have $\forall q \in \mathcal{S}_n$,

$$\|r_n\|_2^2 = (r_n, f - f_n + q - q)_\omega = (r_n, f - q)_\omega + (r_n, q - f_n)_\omega$$
$$= (r_n, f - q)_\omega + (r_n, \tilde{q})_\omega = (r_n, f - q)_\omega \le \|r_n\|_\omega \|f - q\|_\omega$$
$$\therefore \|r_n\|_\omega \le \|f - q\|_\omega$$

□

Optimality

- f_n is a continuous weighted least-squares approximation to f.
- Need $\{\phi_0(x), \phi_1(x), \ldots\}$ a complete set of orthogonal functions in $\mathcal{L}^2_\omega[a, b]$,
- to approximate $f \in \mathcal{L}^2_\omega[a, b]$ over $\mathcal{S}_n$ we must compute the $\alpha_i = (f, \phi_i)_\omega / (\phi_i, \phi_i)_\omega$ and define $f_n(x) = \sum_{i=0}^{n} \alpha_i \phi_i(x)$.
- α_i are typically approximated numerically.
- Note that f_n is easily incremented to f_{n+1} with a single additional coefficient. This is a crucial property in many efficient algorithms that exploit orthogonality, e.g., conjugate directions and conjugate gradient.

Pointwise Error and Convergence

We have convergence in the mean, for $f \in \mathcal{L}^2_\omega[a, b]$

$$\lim_{n \to \infty} \|f - f_n\|_\omega = 0$$

where $f_n(x) = \sum_{i=0}^{n} \alpha_i \phi_i(x)$.

This does not say anything about pointwise error or convergence, i.e.,

$$\lim_{n \to \infty} |f(x) - f_n(x)|$$

Orthogonal Polynomials

Theorem 15.12. *(Christoffel-Darboux) Suppose* $\{\phi_0(x), \phi_1(x), \ldots\}$ *is a complete set of orthogonal polynomials in* $\mathcal{L}^2_\omega[a, b]$ *with* $(\phi_i, \phi_i) = 1$.

$$(x - \xi)G_n(x, \xi) = (x - \xi)\sum_{i=0}^{n} \phi_i(x)\phi_i(\xi)$$

$$= \frac{\alpha_n}{\alpha_{n+1}}\left[\phi_{n+1}(x)\phi_n(\xi) - \phi_{n+1}(\xi)\phi_n(x)\right]$$

where α_n *and* α_{n+1} *are the leading coefficients of* $\phi_n(x)$ *and* $\phi_{n+1}(x)$ *respectively.* $G_n(x, \xi)$ *is called the kernel of the set of orthogonal polynomials.*

Pointwise Error and Convergence

Lemma. *Since it can be shown that*

$$\int_a^b \omega(\xi) G_n(x,\xi) d\xi = 1, \text{ where } G_n(x,\xi) = \sum_{i=0}^{n} \phi_i(x)\phi_i(\xi)$$

and $(\phi_i, \phi_i)_\omega = 1$, *the pointwise error has the form*

$$R_n(x) = f(x) - f_n(x) = f(x) - \sum_{i=0}^{n} \alpha_i \phi_i(x)$$

$$= f(x) - \sum_{i=0}^{n} \phi_i(x) \int_a^b \omega(\xi) f(\xi) \phi_i(\xi) d\xi$$

$$= f(x) - \int_a^b \omega(\xi) f(\xi) \sum_{i=0}^{n} \phi_i(x)\phi_i(\xi) d\xi$$

$$= f(x) - \int_a^b \omega(\xi) G_n(x,\xi) f(\xi) d\xi = \int_a^b \omega(\xi) G_n(x,\xi)(f(x) - f(\xi)) d\xi$$

Pointwise Error and Convergence

Need extra smoothness to state uniform convergence results.

Theorem 15.13. *Suppose* $\{P_0(x), P_1(x), \ldots\}$ *are the Legendre polynomials in* $\mathcal{L}^2_\omega[-1,1]$. *Let* $f \in \mathcal{L}^2_\omega[-1,1]$ *have continuous first and second derivatives. If* $f_n(x)$ *is the optimal polynomial of degree* n *approximating* $f(x)$ *with* $\omega(x) = 1$ *then for any* $\epsilon > 0, \exists n > 0$ *such that* $\forall -1 \leq x \leq 1$

$$|f(x) - f_n(x)| \leq \frac{\epsilon}{\sqrt{n}}$$

$$|f(x) - f_n(x)| = O\big(n^{-1/2}\big)$$

Pointwise Error and Convergence

Theorem 15.14. *Suppose* $\{T_0(x), T_1(x), \ldots\}$ *are the Chebyshev polynomials in* $\mathcal{L}^2_\omega[-1,1]$. *Let* $f \in \mathcal{L}^2_\omega[-1,1]$ *have continuous first and second derivatives. If* $f_n(x)$ *is the optimal polynomial of degree* n *approximating* $f(x)$ *with* $\omega(x) = 1/\sqrt{1-x^2}$ *then for any* $\epsilon > 0, \exists n > 0$ *such that* $\forall -1 \leq x \leq 1$

$$|f(x) - f_n(x)| = O\big(n^{-1}\big)$$

Note. $|f(x) - f_n(x)| = O\big(n^{-1/2}\big)$ can be shown for any of the families of orthogonal polynomials under mild assumptions and a continuous second derivative.

Discrete Least Squares

Suppose $x_0 < x_1 < \cdots < x_m$ are given and the metric

$$\sum_{i=0}^{m} \omega_i (f(x_i) - p_n(x_i))^2$$

with $\omega_i > 0$ is used to determine the polynomial, $p_n^*(x)$, of degree n that achieves the minimal value.

Typically, $m \gg n$. If $m = n$ then the unique interpolating polynomial is the solution.

Discrete Least Squares

Suppose we have a basis of polynomials $\big(\phi_0(x), \ldots, \phi_n(x)\big)$ and let

$$p_n(x) = \sum_{j=0}^{n} \phi_j(x)\xi_j$$

then the conditions are

$$\begin{pmatrix} \rho_0 \\ \rho_1 \\ \vdots \\ \rho_m \end{pmatrix} = \begin{pmatrix} \omega_0^{1/2} f(x_0) \\ \omega_1^{1/2} f(x_1) \\ \vdots \\ \omega_m^{1/2} f(x_m) \end{pmatrix} - \begin{pmatrix} \omega_0^{1/2}\phi_0(x_0) & \ldots & \omega_0^{1/2}\phi_n(x_0) \\ \omega_1^{1/2}\phi_0(x_1) & \ldots & \omega_1^{1/2}\phi_n(x_1) \\ \vdots & & \vdots \\ \omega_m^{1/2}\phi_0(x_m) & \ldots & \omega_m^{1/2}\phi_n(x_m) \end{pmatrix} \begin{pmatrix} \xi_0 \\ \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

$$r = W^{1/2}(b - Ax) = (\tilde{b} - \tilde{A}x)$$

Discrete Least Squares

Two equivalent problems:

$$\min_{x \in \mathbb{R}^n} \|b - Ax\|_W^2$$

$$\text{where } \|v\|_W^2 = v^T W v$$

$$\min_{x \in \mathbb{R}^n} \|\tilde{b} - \tilde{A}x\|_2^2$$

The latter is the standard linear least squares problem. We assume that A has linearly independent columns and therefore we can solve using:

- Householder reflector-based transformation method
- SVD-based method
- Conjugate gradient iterative method

Discrete Least Squares and Polynomials

Back to polynomial approximations.

We assumed a basis of polynomials $\big(\phi_0(x), \dots, \phi_n(x)\big)$ and let

$$p_n(x) = \sum_{j=0}^{n} \phi_j(x)\xi_j$$

Can we select $\phi_j(x)$ so that $\tilde{A} = W^{1/2}A$ has orthogonal columns? If so then the discrete least squares problem is solved by applying $\tilde{A}^T$ to the vector $\tilde{b}$ and scaling.

Discrete Orthogonal Polynomials

Define the inner product

$$(f, g) = (g, f) = \sum_{i=0}^{m} f(x_i)g(x_i)$$

i.e., we have $\omega_i = 1$ for $0 \leq i \leq m$

We want polynomials $P_i(x)$ for $0 \leq i \leq m$ such that

$$(P_r(x), P_s(x)) = \delta_{r,s}, \quad 0 \leq r, s \leq m$$

We restrict the problem further by choosing $-1 \leq x_i \leq 1$ and equally spaced points $x_i = x_0 + ih$, with $x_0 = -1$ and $h = 2/m$.

Gram Polynomials

Theorem 15.15. *Let $m > 0$ be given and let $x_i = x_0 + ih$, with $x_0 = -1$ and $h = 2/m$ and define the inner product*

$$(f, g) = (g, f) = \sum_{i=0}^{m} f(x_i)g(x_i).$$

The Gram polynomials, $P_i(x)$, for $0 \leq n \leq m$, are defined by the recurrence

$$P_{-1}(x) = 0, \quad P_0(x) = \frac{1}{\sqrt{m+1}}, \quad P_{n+1}(x) = \alpha_n x P_n(x) - \gamma_n P_{n-1}(x)$$

$$\alpha_n = \frac{m}{n+1}\left(\frac{4(n+1)^2 - 1}{(m+1)^2 - (n+1)^2}\right)^{1/2} \text{ and } \gamma_n = \frac{\alpha_n}{\alpha_{n-1}}$$

$$\text{satisfy } (P_i, P_j) = \delta_{i,j}, \quad 0 \leq i, j \leq m$$

Proof. See Dahlquist and Bjorck or Isaacson and Keller. □

Gram Polynomials

- Gram Polynomials are the discrete analogs of the Legendre Polynomials
- When $n << \sqrt{m}$ they behave like Legendre polynomials.
- When $n >> \sqrt{m}$ they have large oscillations and large maximum norms.
- When using equidistant data $n < 2\sqrt{m}$ is recommended.

Chebyshev Polynomials

Lemma. *Recall the Chebyshev polynomials*

$$T_n(x) = \cos(n \arccos x), \; n \geq 1$$

$$T_0(x) = 1, \;\; T_1(x) = x, \;\; T_{n+1} = 2xT_n(x) - T_{n-1}(x)$$

and consider the roots of $T_{m+1}(x)$ *for some given* $m > 0$*,*

$$x_i = \cos \frac{2i+1}{m+1} \frac{\pi}{2}, \;\; 0 \leq i \leq m$$

The discrete inner product $(T_i, T_j) = \sum_{k=0}^{m} T_i(x_k) T_j(x_k)$ *for* $0 \leq i, j \leq m$ *satisfies (note the weights are all* 1*):*

$$(T_i, T_j) = 0, \;\; i \neq j$$

$$(T_i, T_j) = \frac{m+1}{2}, \;\; i = j \neq 0$$

$$(T_i, T_j) = m+1, \;\; i = j = 0$$

Chebyshev Polynomials

Theorem 15.16. *Let*

$$P_n(x) = \frac{\sqrt{2}}{\sqrt{m+1}} \cos(n \arccos x), \; n \geq 1$$

$$P_0(x) = \frac{1}{\sqrt{m+1}}, \;\; P_1(x) = \frac{\sqrt{2}}{\sqrt{m+1}} T_1(x),$$

$$P_2(x) = \frac{\sqrt{2}}{\sqrt{m+1}} T_2(x), \quad P_{n+1} = 2xP_n(x) - P_{n-1}(x), \;\; n \geq 2$$

and, for some given $m > 0$, *let*

$$x_i = \cos \frac{2i+1}{m+1} \frac{\pi}{2}, \;\; 0 \leq i \leq m$$

The discrete inner product $(P_i, P_j) = \sum_{k=0}^{m} P_i(x_k) P_j(x_k)$ *for* $0 \leq i, j \leq m$ *satisfies*

$$(P_i, P_j) = \delta_{i,j}$$