ORTHOGONAL POLYNOMIALS

Definition *Let again $X \equiv [a, b] \in R$, and $w : X \to R$ an almost everywhere positive and Riemann integrable function on X. The function w is called* ***weighting function****.*

Definition *Let f and g be two elements of $C[X]$. Given a weighting function w, we can define an* ***inner product*** *on $C[X]$ as:*

$$\langle f, g \rangle \equiv \int_a^b f(x) g(x) w(x) dx$$

Definition *Let* $\Phi \subseteq C[X]$. *The elements of* Φ *are* **mutually orthogonal** *with respect to the weighting function w if and only if:*

$$\langle \phi_k, \phi_j \rangle = \begin{cases} 0 \text{ when } k \neq j \\ \alpha_k > 0 \text{ when } k = j \end{cases}$$

for all $k \neq j$.

Definition *The elements of* Φ *are mutually orthonormal with respect to w if and only if they are mutually orthogonal and* $\alpha_j = 1$ *for all* j.

Theorem (Gram-Schmidt) *Given a weighting function $w(x)$, the sequence of algebraic polynomials $\{Q_j\}_{j=0}^{\infty}$ defined by:*

$$Q_{-1}(x) \equiv 0$$

$$Q_0(x) \equiv 1$$

$$Q_{j+1}(x) \equiv (x - m_j)Q_j(x) - q_j Q_{j-1}(x)$$

for $j \geq 0$, where:

$$m_j \equiv \frac{\langle xQ_j, Q_j \rangle}{\langle Q_j, Q_j \rangle}, \quad q_j \equiv \frac{\langle Q_j, Q_j \rangle}{\langle Q_{j-1}, Q_{j-1} \rangle}$$

is mutually orthogonal with respect to $w(x)$.

- The family $\{Q_n\}$ of algebraic orthogonal polynomials with respect to w is **unique** up to a multiplicative constant.
- More precisely, if $\{P_n\}$ is another family of orthogonal polynomials such that P_n has degree exactly n, then $P_n=\alpha_n Q_n$ for some $\alpha_n \neq 0$.
- Note that $p_n(x)$ and $\alpha p_n(x)$ where $\alpha \neq 0$ share the same zeros.

Corollary *Given a family of algebraic orthogonal polynomials $\{Q_n\}$, each algebraic polynomial can be uniquely written as* $p_n(x) = \sum_{j=1}^{n} \alpha_j Q_j(x)$ *where* $\alpha_j \equiv \langle p_j, Q_j \rangle / \langle Q_j, Q_j \rangle$.

- Thanks to Weierstrass' Theorem, each family of orthogonal polynomials can be considered a basis for the space of all continuous real functions on X.

1. The Chebyshev polynomials T_n are orthogonal on $[-1,+1]$ with respect to the weighting function $w(x) \equiv (1-x^2)^{-\frac{1}{2}}$, since:

$$\int_{-1}^{+1} T_k(x)T_j(x)\frac{dx}{\sqrt{1-x^2}} = \int_0^{\pi} \cos(k\theta)\cos(j\theta)d\theta = \begin{cases} 0, & k \neq j \\ \pi, & k = j = 0 \\ \pi/2, & k = j \neq 0 \end{cases}$$

2. Define $\tilde{T}_0 \equiv 2^{-\frac{1}{2}} T_0$ and $\tilde{T}_n \equiv T_n$ for $n \geq 1$. The $\tilde{T}_n$ are orthonormal on $[-1,+1]$ with respect to the weighting function $w(x) \equiv \frac{2}{\pi}(1-x^2)^{-\frac{1}{2}}$, since:

$$\frac{2}{\pi}\int_{-1}^{+1} \tilde{T}_k(x)\tilde{T}_j(x)\frac{dx}{\sqrt{1-x^2}} = \begin{cases} 0, & k \neq j \\ 1, & k = j = 0 \\ 1, & k = j \neq 0 \end{cases}$$

1. The *Hermite polynomials*, defined as:

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

are orthogonal with respect to $w(x) = e^{-x^2}$ on $[-\infty, +\infty]$ and satisfy the recurrence relation:

$$H_0(x) = 1$$

$$H_1(x) = 2x$$

$$H_{j+1}(x) = 2xH_j(x) - 2jH_{j-1}(x)$$

Note furthermore that:

$$H_j(-x) = (-1)^j H_j(x)$$

Least-squares approximation

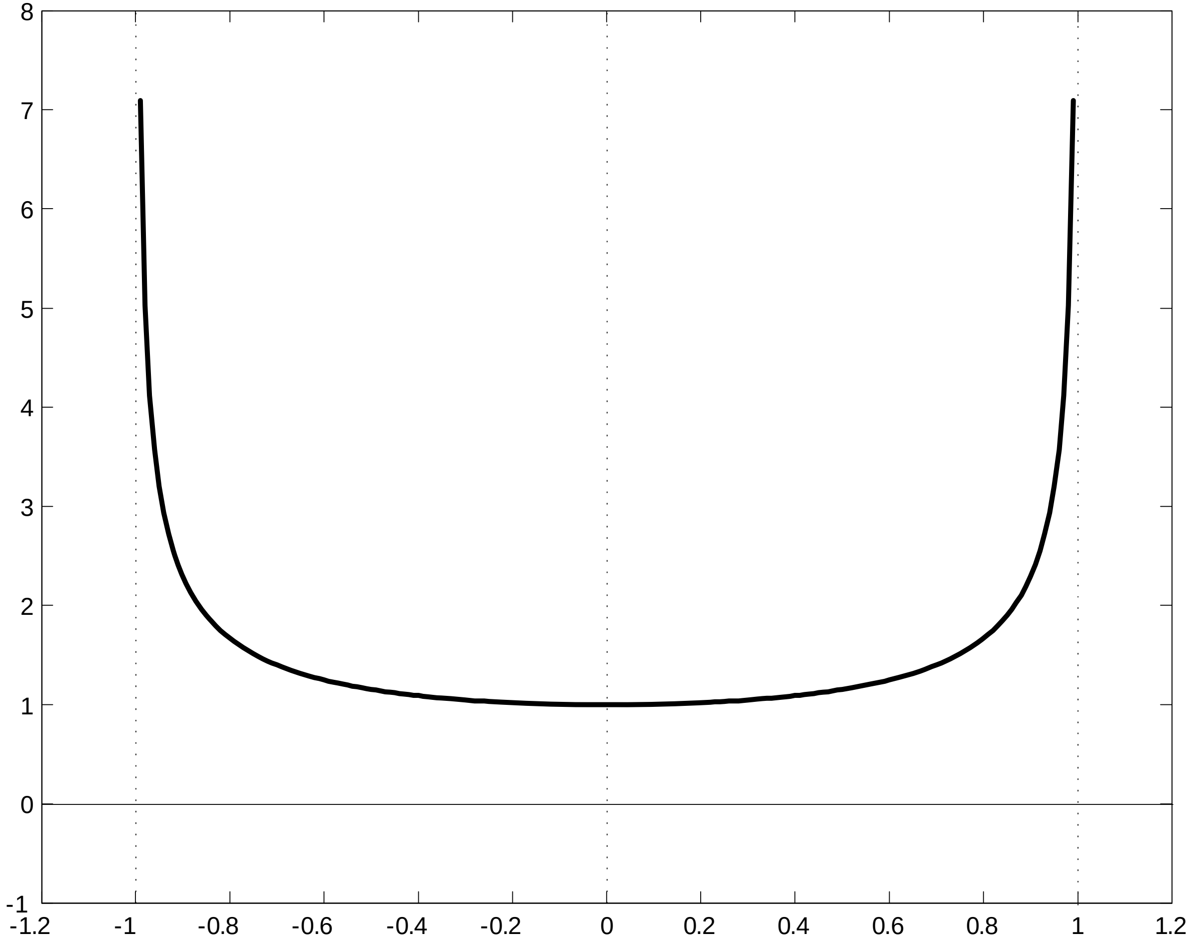
- Let f be an element of $C[X]$ and $w(x)$ a weighting function.
- The function:

$$\|f\|_2 \equiv \sqrt{\int_a^b f(x)^2 w(x)dx} = \sqrt{\langle f,f \rangle}$$

is the L^2*-norm* in $C[X]$, and is strictly convex.

- Note that the sup norm measures the distance between to functions focusing on the "worst" scenario, i.e. the maximum distance in modulus between the two functions, while the L^2-norm is a measure of their "average" distance.

- The weighting function "weights" the squared approximation errors according to x



The weighting function $w_c(x) = (1 - x^2)^{-\frac{1}{2}}$ on $[-1, +1]$.

- Given $\Phi=\{\varphi_j\}$ and *f*, both in *C[X]*, the *n*th degree **least squares polynomial approximation** of *f* w.r.t. a weighting function *w* is the polynomial of degree *n* in the φ_j that solves:

$$\min_{c\in R^n} \int_a^b [f(x) - p(x)]^2 w(x)dx$$

Theorem *Let* Φ *be a finite dimensional distinguished subset of* $C[X]$. *For any function* $f \in C[X]$ *and any weighting function w there is a unique* $\hat{\phi} \in \Phi$ *that solves the problem* $\min_{\phi\in\Phi} \|f - \phi\|_2$.

- First order conditions:

$$\langle f, \phi_j \rangle = \sum_{k=0}^{n} c_k \langle \phi_k, \phi_j \rangle, \quad j = 0, 1, \ldots, n$$

- If Φ is a family of orthogonal polynomials, then the FOCs reduce to:

$$c_j = \frac{\langle f, \phi_j \rangle}{\langle \phi_j, \phi_j \rangle} = \frac{\langle f, \phi_j \rangle}{\alpha_j}, \quad j = 0, 1, \ldots, n$$

Definition *Let* $\Phi = \{T_n\}$ *and* $f \in C[X]$. *The nth degree Chebyshev least square approximation of* f *is* $C_n(f)(x) \equiv \sum_{j=0}^{n} c_j \tilde{T}_j(x)$ *where* $c_0 \equiv \langle f, \tilde{T}_0 \rangle / \pi$ *and* $c_j \equiv \frac{2}{\pi} \langle f, \tilde{T}_j \rangle$ *for* $j = 1, 2, \ldots, n$.

Theorem (Lebesgue) $\| f - C_n(f) \|_\infty \leq 4 \left[1 + \frac{\ln(n+1)}{\pi^2} \right] E_n(f)$.

Corollary $\lim_{n \to \infty} \| f - C_n(f) \|_\infty = 0$.

Theorem *If* $f \in C^k[X]$ *with* $k \geq 2$, *then:*

$$|\hat{c}_j| \leq k \frac{\| f^{(k)} \|_\infty}{j^k}$$

for $j \geq 1$, *and therefore* $\sum_{j=0}^{\infty} |\hat{c}_j| < \infty$.

- In other words, the importance of **high-order monomials** in the *Chebyshev least squares approximation* is rapidly decreasing: the value of $|c_j|$ is falling at the rate $[j/(j+1)]^k - 1$.

LINEAR QUADRATURE

- Let $X \equiv [a,b] \in R$, and let f be a *Riemann integrable* element of $C[X]$.

- Numerical integration, or **quadrature**, is a method to approximate the value of a definite integral like:

$$I(f) = \int_a^b f(x)dx$$

using only linear combinations of values of f:

$$I(f) \approx I_n(f,\omega) \equiv \sum_{j=1}^{n} \omega_j f(x_j)$$

where $\{x_j\}$ are the **quadrature nodes** and $\{\omega_j\}$ the **quadrature weights**.

For the sake of exposition, let us for the moment assume that f is an algebraic polynomial of degree n, i.e. $f(x) = p_n(x) \equiv \sum_{j=0}^{n} c_j x^j$.

For a given set of nodes $\{x_j\}_{j=0}^{n}$, we already know that:

$$p_n(x) = \sum_{j=0}^{n} p_n(x_j) l_{j,n}(x_j)$$

Therefore:

$$\int_a^b p_n(x)dx = \sum_{j=0}^{n} p_n(x_j) \int_a^b l_{j,n}(x)dx$$

where the functions $l_{j,n}(x)$ are the *Lagrange fundamental polynomials* of degree n defined as:

$$l_{j,n}(x) = \prod_{k=0,k\neq j}^{n} \frac{x - x_k}{x_j - x_k}, \quad j = 0, 1, \ldots, n$$

Theorem *Given $n+1$ distinct nodes in X, there exists a unique set of weights $\{\hat{\omega}_j\}_{j=0}^{n}$ such that:*

$$\int_a^b p_m(x)dx = \sum_{j=0}^{n} \hat{\omega}_j p_m(x_j)$$

for all algebraic polynomials of degree $m \leq n$, and:

$$\hat{\omega}_j = \int_a^b l_{j,n}(x)dx$$

More generally, let again f be an integrable element of $C[X]$, and consider $L_n(f)$, the algebraic polynomial of degree n that interpolates f at the given set of nodes $\{x_j\}_{j=0}^{n}$.

Note that:

$$\int_a^b L_n(f)(x)dx = \sum_{j=0}^{n} f(x_j) \int_a^b l_{j,n}(x)dx = \sum_{j=0}^{n} \hat{\omega}_j f(x_j)$$

- Linear quadrature schemes are able, for a given set of $n+1$ nodes, to calculate the exact integral of any algebraic polynomial of degree less or equal to n by choosing the proper weights ω_j.
- If the weights and the nodes are jointly chosen to optimize the accuracy of the approximation, we have further $n+1$ degrees of freedom
- We can therefore expect to calculate the exact integral of algebraic polynomials of degree less or equal to $2n$, which are characterized by $2(n+1)$ coefficients.

Suppose we want to find the weights $\{\omega_0, \omega_1\}$ and the nodes $\{x_0, x_1\}$ that satisfy:

$$\int_{-1}^{+1} p_m(x)dx = \sum_{j=0}^{1} \omega_j p_m(x_j)$$

whenever $m \leq 3$.

Note that:

$$\int_{-1}^{+1} \left(\sum_{j=0}^{3} c_j x^j \right) dx = \sum_{j=0}^{3} c_j \int_{-1}^{+1} x^j dx$$

The monomials $\{x^j\}_{j=0}^{3}$ are algebraic polynomials of degree less or equal 3 themselves.

Hence the solution to our problem is pinned down by the following 4 conditions:

$$\omega_0 1 + \omega_1 1 = \int_{-1}^{+1} 1dx = 2$$

$$\omega_0 x_0 + \omega_1 x_1 = \int_{-1}^{+1} xdx = 0$$

$$\omega_0 x_0^2 + \omega_1 x_1^2 = \int_{-1}^{+1} x^2 dx = \frac{2}{3}$$

$$\omega_0 x_0^3 + \omega_1 x_1^3 = \int_{-1}^{+1} x^3 dx = 0$$

The unique solution to the system is:

$$\omega_0 = 1,\ \omega_1 = 1,\ x_0 = -\frac{\sqrt{3}}{3},\ x_1 = \frac{\sqrt{3}}{3}$$

Hence, the quadrature formula that computes the exact integral of any algebraic polynomial of degree less or equal to 3 using only two quadrature nodes, i.e. the formula of precision 3, is:

$$\int_{-1}^{+1} f(x)dx = f\left(\frac{-\sqrt{3}}{3}\right) + f\left(\frac{\sqrt{3}}{3}\right)$$

Theorem (Gauss) *Let $\{Q_j\}$ be the sequence of orthogonal polynomials relative to a given weighting function $w(x)$, and let $\{x_j\}_{j=1}^{n}$ be the n zeros of Q_n in X. Then, the quadrature formula:*

$$\int_a^b f(x)w(x)dx \approx \sum_{j=1}^{n} \hat{\omega}_j f(x_j)$$

where:

$$\hat{\omega}_j \equiv \int_a^b l_{j,n-1}(x)w(x)dx$$

is exact for all polynomials of degree strictly less than $2n$.

Lemma (Stieltjes) *Let* $\{\hat{\omega}_j\}_{j=1}^{n}$ *be the set of weights defined in (ref: eq10). Then:*

$$\hat{\omega}_j = \langle l_{j,n-1}, l_{j,n-1} \rangle > 0$$

for $j = 1, 2, \ldots, n$.

Theorem (Stieltjes) *The approximation error is bounded in modulus:*

$$|I(f) - I_n(f, \hat{\omega})| \leq 2E_{2n-1}(f) \int_a^b w(x)dx$$

Hence:

$$\lim_{n\to\infty} I_n(f, \hat{\omega}) = I(f)$$

The Golub-Welsch algorithm

Let us now discuss a general method to jointly solve for the quadrature nodes and weights when the orthogonal polynomials $\{Q_n\}$ are characterized by a closed form recurrence formula.

We already know that the Gram-Schmidt procedure iteratively defines the algebraic polynomials that are mutually orthogonal with respect to a particular weighting function.

In general, given the weighting function $w(x)$ and setting $Q_{-1}(x) \equiv 0$ and $Q_0(x) = 1$, we have that:

$$Q_{j+1}(x) = (x - m_j)Q_j(x) - q_j Q_{j-1}(x), \;\; j = 0, 1, 2\ldots$$

For a given degree n, the recurrence relation can be rewritten in matrix form as:

$$x\mathbf{Q} = \mathbf{T} \cdot \mathbf{Q} + Q_n \mathbf{e}_{n-1}$$

where:

$$\mathbf{Q} \equiv \begin{bmatrix} Q_0 \\ Q_1 \\ \vdots \\ Q_{n-1} \end{bmatrix}, \quad \mathbf{e}_{n-1} \equiv \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}, \quad T \equiv \begin{bmatrix} a_0 & 1 & & & & \\ b_1 & a_1 & 1 & & & \\ & & \vdots & \ddots & & \\ & & b_{n-2} & a_{n-2} & 1 \\ & & & b_{n-1} & a_{n-1} \end{bmatrix}$$

To simplify the task, note that eigenvalues are preserved by similarity transformations.

Note that if x_j is an eigenvalue of $\mathbf{T}$, then $Q_n(x_j) = 0$. Hence finding the eigenvalues of $\mathbf{T}$ is equivalent to finding the roots of Q_n.

We can apply a diagonal similarity transformation **D** to **T** obtaining:

$$\mathbf{J} \equiv \mathbf{D}\mathbf{T}\mathbf{D}^{-1} = \begin{bmatrix} a_0 & \sqrt{b_1} & & & \\ \sqrt{b_1} & a_1 & \sqrt{b_2} & & \\ & & \vdots & \ddots & \\ & & \sqrt{b_{n-2}} & a_{n-2} & \sqrt{b_{n-1}} \\ & & & \sqrt{b_{n-1}} & a_{n-1} \end{bmatrix}$$

The Jacobian matrix **J** is a symmetric tridiagonal matrix whose eigenvalues coincide with the eigenvalues of **T**, and therefore with the zeros of Q_n.

Furthermore, if $\mathbf{v}_j$ is the eigenvector (normalized so that $\mathbf{v} \cdot \mathbf{v} = 1$) associated to the eigenvalue x_j, then:

$$w_j = v_{j,1}^2 \int_a^b w(x)dx$$

where $v_{j,1}$ is the first element of $\mathbf{v}_j$.

Gauss-Chebyshev quadrature

- The Chebyshev polynomials are orthogonal with respect to $w(x) \equiv (1-x^2)^{-1/2}$ over $X \equiv [-1,+1]$.
- The Gauss-Chebyshev weights happen to be constant and equal to π/n.
- Therefore, the **Gauss-Chebyshev quadrature formula** is simply the following:

$$\int_{-1}^{+1} \frac{f(x)}{\sqrt{1-x^2}} dx \approx \frac{\pi}{n} \sum_{j=1}^{n} f(x_j)$$

where:

$$x_j = \cos\left(\frac{2j-1}{2n} \pi \right), \quad j = 1, 2, \ldots n$$

- Using the change of variable $x=(y+1)(b-a)/2+a$ we can write:

$$\int_a^b f(x)dx = \frac{b-a}{2}\int_{-1}^{+1} f\left[\frac{(y+1)(b-a)}{2}+a\right]\frac{\sqrt{1-y^2}}{\sqrt{1-y^2}}dy$$

- Hence:

$$\int_a^b f(x)dx \approx \frac{b-a}{2}\frac{\pi}{n}\sum_{j=1}^{n} f\left[\frac{(y_j+1)(b-a)}{2}+a\right]\sqrt{1-y_j^2}$$

Gauss-Hermite quadrature

- The Hermite polynomials are orthogonal with respect to $w(x)=e^{-x^2}$ over $X\equiv[-\infty,+\infty]$.
- The Gauss-Hermite quadrature formula is:

$$\int_{-\infty}^{+\infty} f(x)e^{-x^2}dx \approx \sum_{j=1}^{n} \omega_j f(x_j)$$

where the ω_j are the Hermite weights and the x_j are the n zeros of $H_n(x)$.

- Recall that, if $x\sim N(\mu,\sigma^2)$, then:

$$E[f(x)] = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty} f(x)e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

- Using the change of variable $y=(x-\mu)/(\sqrt{2}\sigma)$ we can rewrite it as:

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty} f(x)e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty} f\left(\sqrt{2}\,\sigma y+\mu\right)e^{-y^2}dy$$

- Hence, $E[f(x)]$ can be approximated by:

$$E[f(x)] \approx \frac{1}{\sqrt{\pi}}\sum_{j=1}^{n}\omega_j f\left(\sqrt{2}\,\sigma y_j+\mu\right)$$

where the y_j are the Gauss-Hermite quadrature nodes.