Approximation and Quadrature

1 Approximation using orthogonal projection

Let $V$ be a vector space with inner product $(u, v)$ and norm $\|u\| = (u, u)^{1/2}$.

We are given $u \in V$ and a subspace $\tilde{V} = \text{span} \{v^{(1)}, \ldots, v^{(n)}\}$ where $v^{(1)}, \ldots, v^{(n)}$ are linearly independent. We want to find $p \in \tilde{V}$ such that $\|u - p\|$ is minimal.

Theorem 1. There is a unique $p \in \tilde{V}$ minimizing $\|u - p\|$. It satisfies the “normal equations”: $u - p \perp \tilde{V}$, i.e.

$$(u - p, v^{(j)}) = 0, \quad j = 1, \ldots, n$$

Method 1 to find $p$: With $p = c_1v^{(1)} + \cdots + c_nv^{(n)}$ we obtain an $n \times n$ linear system:

$$Mc = b, \quad M_{jk} = (v^{(k)}, v^{(j)}) \quad b_j = (u, v^{(j)})$$

Method 2 to find $p$: Step 1: use Gram-Schmidt orthonormalization to find an orthogonal basis $p^{(1)}, \ldots, p^{(n)}$ for $\tilde{V}$:

$$p^{(1)} := v^{(1)}, \quad p^{(2)} := v^{(2)} - \frac{(v^{(2)}, p^{(1)})}{(p^{(1)}, p^{(1)})} p^{(1)}, \quad p^{(3)} := v^{(3)} - \frac{(v^{(3)}, p^{(1)})}{(p^{(1)}, p^{(1)})} p^{(1)} - \frac{(v^{(3)}, p^{(2)})}{(p^{(2)}, p^{(2)})} p^{(2)}, \ldots$$

Define the upper triangular matrix $S = \begin{bmatrix} 1 & s_{12} & \cdots & s_{1n} \\ \vdots & \ddots & \vdots & \vdots \\ \vdots & \ddots & 1 \\ 1 & \end{bmatrix}$. Considering $v^{(j)}, p^{(j)}$ as “column vectors” we can write $[v^{(1)}, \ldots, v^{(n)}] = [p^{(1)}, \ldots, p^{(n)}] S$ and

$$p = c_1v^{(1)} + \cdots + c_nv^{(n)} = [v^{(1)}, \ldots, v^{(n)}] c$$

so the coefficient vectors $c$ and $d$ are related by $Sc = d$.

Step 2: With $p = d_1p^{(1)} + \cdots + d_n p^{(n)}$ the normal equations give $d_j = \frac{(u, p^{(j)})}{(p^{(j)}, p^{(j)})}$.

If we need the coefficient vector $c$ we can solve $Sc = d$ by back substitution.

Application 1: Let $V = \mathbb{R}^N$ then we are given a matrix $A := [v^{(1)}, \ldots, v^{(n)}] \in \mathbb{R}^{N \times n}$. Gram-Schmidt gives a decomposition $Q = PS$. Dividing the columns of $P$ by $\|p^{(j)}\|$ and multiplying the rows or $S$ by $\|p^{(j)}\|$ gives $A = QR$ where $Q$ has orthonormal columns, i.e., $Q^\top Q = I$. We can then solve the least squares problem $\|Ac - u\|_2 = \min$ by $d := Q^\top u$ and solving $Sc = d$. 
**Application 2:** Let $V = L^2_w([a, b])$ which is the space of functions on the interval $[a, b]$ such that $\int_a^b u(x)^2 w(x)dx < \infty$, with a weight function $w(x) \geq 0$. Here $(u, v) := \int_a^b u(x)v(x)w(x)dx$. We consider the subspace $P_n := \text{span} \{1, x, \ldots, x^n\}$. Applying Gram-Schmidt to $1, x, x^2, \ldots$ gives orthogonal polynomials $p_0, p_1, p_2, \ldots$ where $p_j(x) = x^j$ + lower order terms.

**Example 1:** For the interval $[-1, 1]$ and $w(x) = 1$ we obtain the Legendre polynomials.

**Example 2:** For the interval $[-1, 1]$ and $w(x) = (1 - x^2)^{-1/2}$ we obtain the Chebyshev polynomials.

We use a different normalization: With $T_0(x) := 1$ and $T_1(x) := x$ we define

$$T_{n+1}(x) := 2x \cdot T_n(x) - T_{n-1}(x)$$

and have with the change of variables $x = \cos t$

$$T_n(x) = \cos(nt)$$

Hence the zeros $\tilde{x}_j$ and extrema $x_j$ on $[-1, 1]$ are given by

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

with $T_n(x_j) = (-1)^{j-1}$. Note that

$$(T_k, T_\ell) = \int_{t=0}^{\pi} \cos(kt) \cos(\ell t) dt = \begin{cases} 0 & \text{for } k \neq \ell \\ \pi & \text{for } k = \ell = 0 \\ \frac{\pi}{2} & \text{for } k = \ell > 0 \end{cases} \quad (2)$$

## 2 Approximation of functions by polynomials

We consider a continuous function $u(x)$ on the interval $[-1, 1]$ and want to find a polynomial $p_n \in P_n$ such that

$$\|u - p_n\|_{\infty} = \max_{x \in [-1, 1]} |u(x) - p_n(x)|$$

is small. The **Weierstrass approximation theorem** states that we can make the approximation error arbitrarily small:

**Theorem 2.** If $u(x)$ is continuous on $[-1, 1]$ there exists a sequence $p_n \in P_n, n = 0, 1, 2, \ldots$ such that

$$\|u - p_n\|_{\infty} \to 0 \quad \text{as} \quad n \to \infty$$

**Method 1: Interpolation.** We pick $n + 1$ distinct points $x_1, \ldots, x_{n+1} \in [-1, 1]$ . Then there is a unique interpolating polynomial $p_n \in P_n$ satisfying $p_n(x_j) = f(x_j)$ for $j = 1, \ldots, n + 1$. We can write it in terms of the Lagrange formula:

$$p_n(x) = u(x_1)\ell_1(x) + \cdots + u(x_{n+1})\ell_{n+1}(x) \quad \text{where} \quad \ell_j(x) := \prod_{k=1 \ldots n+1}^{k \neq j} \frac{x - x_k}{x_j - x_k}$$

Note that for equidistant nodes the error can be very large. From now on we will use for $x_j$ the extrema of $T_n(x)$, see (1).
Method 2: Orthogonal projection with weight function $w(x) = (1 - x^2)^{-1/2}$. We define $u_n \in \mathcal{P}_n$ as the polynomial which minimizes $\|u - u_n\|_{L_w^2}$. Using (2) we obtain

$$u_n = a_0 T_0(x) + \cdots + a_n T_n(x), \quad a_k = \frac{(u, T_k)}{(T_k, T_k)} = \begin{cases} \frac{1}{\pi} (u, T_k) & \text{for } k = 0 \\ \frac{2}{\pi} (u, T_k) & \text{for } k > 0 \end{cases}$$

Note that $(u, T_k) = \int_{-1}^{1} u(x)T_k(x)(1 - x^2)^{-1/2}dx = \int_{t=0}^{\pi} u(\cos t) \cos(nt)dt$.

For any $u \in L_w^2$ we have that $\|u - u_n\|_{L_w^2} \to 0$ as $n \to \infty$. If $u$ is continuous we may not have pointwise convergence $u_n(x) \to u(x)$ as $n \to \infty$ in a point $x$. With a stronger assumption on $u$ one can prove nicer convergence properties:

**Theorem 3.** Assume that $u(x)$ satisfies a Lipschitz condition $|u(x) - u(\tilde{x})| \leq L |x - \tilde{x}|$ for $x, \tilde{x} \in [-1,1]$. Then the Chebyshev series $u = \sum_{k=0}^{\infty} a_k T_k$ converges absolutely for each $x \in [-1,1]$ and we have uniform convergence

$$\|u - u_n\|_\infty \to 0 \quad \text{as } n \to \infty$$

We can now get a bound for the approximation error for the projection $u_n$ and the interpolation $p_n$ in terms of the Chebyshev coefficients $a_k$:

$$\|u - u_n\|_\infty \leq \sum_{j=n+1}^{\infty} |a_j|$$

$$\|u - p_n\|_\infty \leq 2 \sum_{j=n+1}^{\infty} |a_j|$$

The decay of the coefficients $a_k$ depends on the “smoothness” of the function $u(x)$.

The **total variation** $TV(f)$ of a function is defined as a supremum over all partitions $a = z_0 < z_1 < \cdots < z_m = b$ of the interval $[a, b]$:

$$TV(f) = \sup_{\text{partitions}} \sum_{j=1}^{m} |f(z_j) - f(z_{j-1})|$$

If the function $f$ is piecewise continuously differentiable on a partition $a = X_0 < X_1 < \cdots < X_M = b$ with jumps we have

$$TV(f) = \sum_{j=1}^{M} \int_{X_{m-1}}^{X_m} |f'(x)| \, dx + \sum_{j=1}^{M-1} |f(X_j + 0) - f(X_j - 0)|$$

E.g., the total variation of $\text{sign}(x)$ on $[-1,1]$ is 2. Note that we can formally write $TV(f)$ as $\int_{a}^{b} |f'(x)| \, dx$ if we understand $f'(x)$ as a generalized function with delta functions at the jumps.

**Theorem 4.** If the derivative $u^{(\nu)}(x)$ has bounded total variation $TV(u^{(\nu)})$ we have for $\nu \geq 0$ and $k > \nu$

$$|a_k| \leq \frac{2}{\pi} \frac{TV(u^{(\nu)})}{k(k-1) \cdots (k-\nu)} = O\left(k^{-\nu-1}\right)$$
E.g., for \( f(x) = |x| \) we have \( \nu = 1 \) since \( f'(x) = \text{sign}(x) \) has bounded variation: \( TV(f') = 2 \).

Therefore we obtain for the error of \( u_n \) and \( p_n \) for \( \nu \geq 1 \) and \( n \geq \nu \)

\[
\|u - u_n\|_\infty \leq \frac{2}{\pi} \cdot \frac{1}{\nu} \cdot \frac{TV(u^{(\nu)})}{n(n-1) \cdots (n-\nu+1)} = O\left(n^{-\nu}\right) \tag{3}
\]

\[
\|u - p_n\|_\infty \leq \frac{4}{\pi} \cdot \frac{1}{\nu} \cdot \frac{TV(u^{(\nu)})}{n(n-1) \cdots (n-\nu+1)} = O\left(n^{-\nu}\right)
\]

Note that for \( u(x) = |x| \) we obtain \( \|u - u_n\|_\infty = O(n^{-1}) \) and \( \|u - p_n\|_\infty = O(n^{-1}) \).

We now consider a function \( u(x) \) which is \textbf{analytic} on \([-1, 1]\), i.e., for each \( x_0 \in [-1, 1] \) the Taylor series about \( x_0 \) converges for \( |x - x_0| \leq \epsilon_{x_0} \) with \( \epsilon_{x_0} > 0 \). This implies that the Taylor series actually converges in the complex plane for \( |z - x_0| \leq \epsilon_{x_0} \). Therefore the function \( u(x) \) has an analytic continuation to a function \( u(z) \) in the complex plane, defined in a region containing the interval \([-1, 1]\). Such a function can be extended in the complex plane until we hit singularities (like e.g. poles). The size of the region in the complex plane where \( u(z) \) is analytic determines the decay of \( a_k \).

Let \( \rho > 1 \). The mapping \( w = \frac{1}{2} (z + z^{-1}) \) maps the circle \( |z| = \rho \) to an ellipse \( E_\rho \). This ellipse has foci \(-1, 1\) and the sum of the semiaxes \( a, b \) is \( a + b = \rho \). Note that the circle \( |z| = \rho^{-1} \) is mapped to the same ellipse.

We define \( E_\rho \) as the open region bounded by the curve \( C_\rho \).

**Theorem 5.** Assume that \( u(x) \) is analytic on \([-1, 1]\) and has an analytic extension to the open region \( E_\rho \) with \( \rho > 1 \). Let \( |u(z)| \leq M \) for \( z \in E_\rho \). Then

\[
|a_k| \leq 2M\rho^{-k}
\]

Therefore we obtain for the error of \( u_n \) and \( p_n \) for \( \nu \geq 1 \)

\[
\|u - u_n\|_\infty \leq \frac{2M}{\rho - 1}\rho^{-n} \tag{4}
\]

\[
\|u - p_n\|_\infty \leq \frac{4M}{\rho - 1}\rho^{-n}
\]

**Example:** The function \( u(x) = (1 + x^2)^{-1} \) is analytic on \([-1, 1]\). The function \( u(z) = (1 + z^2)^{-1} \) is analytic everywhere in the complex plane except the two points \( z = \pm i \). We want to find the ellipse \( C_\rho \) which passes through the point \( i \): We must have that the point \( z = i\rho \) on the circle gets mapped to the point \( w = \frac{1}{2}(z + z^{-1}) = \frac{1}{2}(i\rho - (i\rho)^{-1}) = i\rho (\rho - 1)^{-1} = i \cdot 1 \). Hence we need to find \( \rho > 1 \) such that \( \frac{1}{2}(\rho - 1)^{-1} = 1 \). This gives a quadratic equation for \( \rho \) which we can solve:

\[
\rho^2 - 2\rho - 1 = 0, \quad \rho = \frac{2 + \sqrt{8}}{2} = 1 + \sqrt{2}
\]

(we use “+” since we want to find \( \rho > 1 \)). Let \( \rho_* = 1 + \sqrt{2} \). For any \( \rho < \rho_* \) we have that \( u(z) \) is bounded on \( C_\rho \). Hence we obtain that for any \( \rho < \rho_* = 1 + \sqrt{2} \) we have

\[
\|u - u_n\|_\infty \leq C_\rho \rho^{-n}, \quad \|u - p_n\|_\infty \leq 2C_\rho \rho^{-n} \tag{5}
\]
3 Quadrature

For an interval \([a, b]\) and a “weight function” \(w(x) \geq 0\) we want to approximate integrals of the form

\[
I(f) := \int_a^b f(x)w(x)dx
\]

We pick distinct nodes \(x_1, \ldots, x_n \in [a, b]\). Let \(p(x) = f(x_1)\ell_1(x) + \cdots + f(x_n)\ell_n(x)\) denote the interpolating polynomial, then we define the approximation as

\[
Q(f) = \int_a^b p(x)w(x)dx = \sum_{j=1}^n w_j f(x_j)
\]

where \(w_j = I(\ell_j)\). Note that this quadrature rule is exact if \(f \in P_{n-1}\), i.e., \(I(f) = Q(f)\).

Usually it is easier to determine \(w_1, \ldots, w_n\) by using the \(n\) equations \(Q(x^j) = I(x^j)\) for \(j = 0, \ldots, n-1\). This is a linear system of \(n\) equations for \(n\) unknowns.

Example: For \(w(x) = 1\) on \([-1, 1]\) and the nodes \(x_1 = -1, x_2 = 0, x_3 = 1\) we obtain for \(f(x) = 1, x, x^2\) the three equations

\[
\begin{align*}
w_1 \cdot 1 + w_2 \cdot 1 + w_3 \cdot 1 &= \int_{-1}^{1} 1dx = 2 \\
w_1 \cdot (-1) + w_2 \cdot 0 + w_3 \cdot 1 &= \int_{-1}^{1} x dx = 0 \\
w_1 \cdot 1 + w_2 \cdot 0 + w_3 \cdot 1 &= \int_{-1}^{1} x^2 dx = \frac{2}{3}
\end{align*}
\]

This linear system has the solution \(w_1 = \frac{1}{3}, w_2 = \frac{4}{3}, w_3 = \frac{1}{3}\) which is the well known Simpson rule.

Note that the quadrature rule \(Q(f)\) may be exact for polynomials \(p \in P_m\) with \(m > n - 1\). But it can never be exact for \(P_{2m}\) since for \(f(x) = (x - x_1)^2 \cdots (x - x_n)^2\) we have \(Q(f) = 0\) and \(I(f) > 0\).

It turns out that we can pick nodes \(x_1, \ldots, x_n\) such that the quadrature rule is exact for \(p \in P_{2n-1}\), this is the so-called **Gaussian quadrature**:

**Theorem 6.** Let \(p(x)\) denote the orthogonal polynomial of degree \(n\) for the weight function \(w(x)\) on \([a, b]\). Let \(x_1, \ldots, x_n\) be the zeros of the function \(p(x)\), and let \(w_1, \ldots, w_n\) denote the corresponding coefficients of the quadrature formula. Then

- the nodes \(x_1, \ldots, x_n\) are distinct points in \((a, b)\)
- \(w_j > 0\) for \(j = 1, \ldots, n\)
- \(Q(p) = I(p)\) for all \(p \in P_{2n-1}\)

**Example:** Approximate \(I = \int_0^1 x^{1/3} \cos x \, dx\) by a Gaussian rule with 1 node.

If we use \(f(x) = x^{1/3} \cos x\) and \(w(x) = 1\) the function \(f\) is not very smooth and we will only have very slow convergence of \(Q_n(f)\).
Therefore we let \( f(x) = \cos x \) and \( w(x) = x^{1/3} \). For the Gauss rule we first need to find the orthogonal polynomial of degree 1: \( p_1(x) = x - c \) should be orthogonal on 1, i.e.,

\[
0 = \int_0^1 (x - c) \cdot 1 \cdot x^{1/3} \, dx = \left[ \frac{3}{7} x^{7/3} - \frac{3}{4} x^{4/3} \right]_{x=0}^{1} = \frac{3}{7} - \frac{3}{4}
\]

to find \( c = \frac{4}{7} \). Now \( x_1 \) is the zero of \( p_1(x) = x - \frac{4}{7} \), so \( x_1 = \frac{4}{7} \). We need to determine \( w_1 \) such that \( Q(1) = I(1) \), i.e.,

\[
w_1 \cdot 1 = \int_0^1 1 \cdot x^{1/3} \, dx = \left[ \frac{3}{4} x^{4/3} \right]_{x=0}^{1} = \frac{3}{4}
\]

hence \( w_1 = \frac{3}{4} \). Therefore we have obtain the Gaussian rule

\[
\int_0^1 f(x) x^{1/3} \, dx \approx \frac{3}{4} \cdot f\left(\frac{4}{7}\right)
\]

and we obtain for \( f(x) = \cos x \) the approximation \( Q(f) = \frac{3}{4} \cos\left(\frac{4}{7}\right) = 0.6308 \). The exact value is \( I(f) = 0.6076 \).

**Quadrature error.** The quadrature error satisfies

\[
I(f) - Q(f) = \int_a^b (f(x) - p(x)) \, w(x) \, dx
\]

where \( p(x) \) is the interpolating polynomial through the nodes \( x_1, \ldots, x_n \).

For many quadrature rules we have that \( w_j \geq 0 \) for \( j = 1, \ldots, n \), e.g.,

- any Gaussian quadrature rule
- using Chebyshev nodes for \([-1, 1] \) with \( w(x) = 1 \)

We can then express the quadrature error in terms of any approximating polynomial \( q(x) \):

**Theorem 7.** Assume that the quadrature rule \( Q(f) = w_1 f(x_1) + \cdots + w_n f(x_n) \) satisfies \( w_j \geq 0 \) for \( j = 1, \ldots, n \) and is exact for all polynomials \( p \in \mathcal{P}_m \). Then we have for any \( q \in \mathcal{P}_m \)

\[
|I(f) - Q(f)| \leq 2I(1) \|f - q\|_{\infty}
\]

Consider the interval \([-1, 1] \). Then we can use for \( q(x) \) the Chebyshev projection \( u_m \in \mathcal{P}_m \) and our approximation results (3), (4) to obtain bounds for the quadrature error.

**Example:** We consider Gaussian quadrature for \([-1, 1] \) and \( w(x) = 1 \) with \( n \) nodes. Then we have \( m = 2n - 1 \) and \( I(1) = \int_{-1}^{1} 1 \, dx = 2 \).

(i) Let \( f(x) = |x| \). Then we have \( \nu = 1 \) and (6), (3) give

\[
|Q_n(f) - I(f)| \leq 2 \cdot \frac{2}{\pi} \cdot \frac{1}{\nu} \cdot \frac{TV(f')}{m} = \frac{8}{\pi} \cdot \frac{2}{2n - 1}
\]

This is not a sharp estimate, there actually holds \( |Q(f) - I(f)| \leq Cn^{-2} \).

(ii) Let \( f(x) = \frac{1}{1 + x^2} \). Then \( f(z) \) is analytic in \( E_{\rho} \) for \( \rho < 1 + \sqrt{2} \). Hence we have for any \( \rho < 1 + \sqrt{2} \) using (6), (5)

\[
|Q_n(f) - I(f)| \leq 2 \cdot \frac{2}{\pi} \cdot C_{\rho} \nu^{-(2n-1)} = C_{\rho} e^{-an} \quad \text{with} \quad a = 2 \log(1 + \sqrt{2}) = 1.763
\]

Here the estimate for \( a \) is very sharp: Plotting \( \log|Q_n - I| \) vs. \( n \) gives a straight line with slope \( -1.763 \).