Final

MATH 104A Intro to Numerical Analysis

Harry Coleman

December 13, 2020

1

In this problem, we want to approximate the definite integral

$$\int_0^{\frac{\pi}{2}} \sin 2x \, dx$$

using various quadrature rules.

1.1

Using the **Trapezoidal Rule** to approximate the integral.

The trapezoidal approximation of this integral gives us

$$\int_0^{\frac{\pi}{2}} \sin 2x \, dx \approx \frac{1}{2} (\sin \pi + \sin 0) \left(\frac{\pi}{2} - 0 \right) = \frac{\pi}{4}(0 - 0) = 0.$$

1.2

Using the **Composite Trapezoidal Rule** and the grid $x_j = jh$ with $h = \frac{\pi}{24}$ and $j = 0, 1, \ldots, 12$ to approximate the integral.

The composite trapezoidal approximation of the integral of a function f on the interval $[a, b]$, using the equidistributed nodes $x_j = a + jh$ for $j = 0, \ldots, n$ is given by

$$\int_a^b f(x)\,dx \approx \frac{h}{2}\left(f(a) + \sum_{j=1}^{n-1} f(a + jh) + f(b)\right).$$

We approximate the integral of $f(x) = \sin 2x$ on the interval $[0, \pi/2]$, taking $h = \frac{\pi}{24}$ and $n = 12$. The following code evaluates the above formula for these parameters.

```
import math

def CTR():
    h = math.pi/24
    n = 12
    F = [math.sin(2*j*h) for j in range(0, n+1)]

    print((F[0] + 2*sum(F[1:n]) + F[-1]) * h/2)

CTR()
```

Running this code (see Jupyter file) gives us the approximation

$$\int_0^{\frac{\pi}{2}} \sin 2x\,dx \approx 0.9942818882921578.$$

1.3

Using the **Composite Simpson's Rule** and the grid $x_j = jh$ with $h = \frac{\pi}{24}$ and $j = 0, 1, 2, \ldots, 12$ to approximate the integral.

The composite Simpson's rule approximation of the integral of a function f on the interval $[a, b]$, using the equidistributed nodes $x_j = a + jh$ for $j = 0, \ldots, n$ is given by

$$\int_a^b f(x)\,dx \approx \frac{h}{3}\left(f(x_0) + f(x_n) + 2\sum_{j=2}^{n/2-1} f(x_{2j}) + 4\sum_{j=1}^{n/2} f(x_{2j-1})\right).$$

We approximate the integral of $f(x) = \sin 2x$ on the interval $[0, \pi/2]$, taking $h = \frac{\pi}{24}$ and $n = 12$. The following code evaluates the above formula for these parameters.

```
def CSR():
    h = math.pi/24
    n = 12
    F = [math.sin(2*j*h) for j in range(0, n+1)]

    print((F[0] + F[-1] + 2*sum(F[2:n:2]) + 4*sum(F[1:n:2])) * h/3)

CSR()
```

Running this code (see Jupyter file) gives us the approximation

$$\int_0^{\frac{\pi}{2}} \sin 2x\,dx \approx 1.0000263121705928.$$

1.4

Let the uniform grid $x_j = jh$ with $h = \frac{\pi}{2n}$ and $j = 0, 1, \ldots, n$ for a **positive integer** n. Using the formula for error term estimate of **Composite Trapezoidal Rule** to determine values of n that ensure an approximation error less than 0.00002. You are allowed to use a calculator to compute n.

The error of the composite trapezoidal rule on these parameters is given by

$$\frac{1}{12}\left(\frac{\pi}{2} - 0\right)\left(\frac{\pi}{2n}\right)^2 \sin 2\xi = \frac{\pi^3 \sin 2\xi}{96n^2}, \quad \xi \in (0, \pi/2).$$

Note that the maximum $\sin 2\xi$ attains on $(0, \pi/2)$ is 1, for $\xi = \pi/4$. Then the error is bounded by

$$\frac{\pi^3}{96n^2}.$$

Bounding this value for a given $\varepsilon > 0$, we find

$$\frac{\pi^3}{96n^2} < \varepsilon,$$

$$\frac{\pi^3}{96\varepsilon} < n^2,$$

$$\sqrt{\frac{\pi^3}{96\varepsilon}} < n.$$

Taking $\varepsilon = 0.00002$ (see Jupyter file), we obtain a lower bound

$$n > \sqrt{\frac{\pi^3}{96 \cdot 0.00002}} \approx 127.07911881051172.$$

Thus, for $n \geq 128$, we ensure an approximation error less than 0.00002.

1.5

Suppose a quadrature rule $Q_h(f)$ for $\int_a^b f(x)\,dx$ satisfies

$$Q_h(f) = \int_a^b f(x)\,dx + c_1 h^4 + c_2 h^6 + \cdots + c_k h^{2k+2} + \cdots .$$

Find a formula that combines $Q_h(f)$ and $Q_{\frac{h}{2}}(f)$ to give $\int_a^b f(x)\,dx + O(h^6)$.

Let I denote the integral

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

Using the formulas for $Q_h(f)$ and $Q_{\frac{h}{2}}(f)$ we have the following:

$$I = Q_h(f) - c_1 h^4 - \sum_{k=2}^{\infty} c_k h^{2k+2}$$

$$I = Q_{\frac{h}{2}}(f) - c_1 \frac{h^4}{2^4} - \sum_{k=2}^{\infty} c_k \frac{h^{2k+2}}{2^{2k+2}}.$$

We combine these to cancel the h^4 terms, obtaining

$$I - 16I = Q_h(f) - 16Q_{\frac{h}{2}}(f) - \sum_{k=2}^{\infty} c_k h^{2k+2} + 16\sum_{k=2}^{\infty} c_k \frac{h^{2k+2}}{2^{2k+2}},$$

$$-15I = Q_h(f) - 16Q_{\frac{h}{2}}(f) - \sum_{k=2}^{\infty} c_k \left(1 - \frac{1}{2^{2k-2}}\right) h^{2k+2}.$$

Thus, we obtain the combined quadrature rule

$$\frac{Q_h(f) - 16Q_{\frac{h}{2}}(f)}{15} = \int_a^b f(x)\,dx + d_2 h^6 + \cdots + d_k h^{2k+2} + \cdots ,$$

where

$$d_k = \frac{c_k}{15}\left(1 - \frac{1}{2^{2k-2}}\right).$$

2

Let $s_0 = 1 + (x+1)^3$ for $-1 \leq x \leq 0$. Determine $s_1(x)$ for $0 \leq x \leq 1$ such that

$$s(x) = \begin{cases} s_0(x) & \text{for } -1 \leq x \leq 0 \\ s_1(x) & \text{for } \quad 0 \leq x \leq 1 \end{cases}$$

is a natural cubic spline on $[-1, 1]$ with nodes at $-1, 0, 1$.

Let the cubic polynomial $s_1(x)$ be given by

$$s_1(x) = ax^3 + bx^2 + cx + d.$$

In order for $s(x)$ to be a natural cubic spline, the following must be satisfied

$$\begin{aligned} s_1(0) &= s_0(0), \\ s_1'(0) &= s_0'(0), \\ s_1''(0) &= s_0''(0), \\ s_1''(1) &= 0. \end{aligned}$$

We now find

$$\begin{aligned} s_1(0) &= a(0)^3 + b(0)^2 + c(0) + d = d, \\ s_0(0) &= 1 + (0+1)^3 = 2, \end{aligned}$$

and the first condition implies $d = 2$. Next,

$$\begin{aligned} s_1'(0) &= 3a(0)^2 + 2b(0) + c = c, \\ s_0'(0) &= 3(0+1)^2 = 3, \end{aligned}$$

and the second condition implies $c = 3$. Next,

$$\begin{aligned} s_1''(0) &= 6a(0) + 2b = 2b, \\ s_0''(0) &= 6(0+1) = 6, \end{aligned}$$

and the third condition implies $b = 3$. Finally,

$$s_1''(1) = 6a(1) + 2b = 6a + 6,$$

and the fourth condition implies $a = -1$. Thus, we have

$$s_1(x) = -x^3 + 3x^2 + 3x + 2.$$

3

Given two discrete, periodic functions $\{f_m\}_{m=0}^{N-1}$ and $\{g_m\}_{m=0}^{N-1}$, we define their convolution, $f * g$, as the discrete function given by

$$(f * g)_m = \sum_{s=0}^{N-1} f_{m-s} g_s, \quad m = 0, 1, \ldots, N-1.$$

We define the Discrete Fourier Transform of $\{f_m\}_{m=0}^{N-1}$ by

$$\widehat{f}_k = \sum_{m=0}^{N-1} f_m e^{-\frac{2\pi i m k}{N}}, \quad k = 0, 1, \ldots, N-1.$$

3.1

Show that

$$f_m = \frac{1}{N} \sum_{k=0}^{N-1} \widehat{f}_k e^{\frac{2\pi i m k}{N}}, \quad m = 0, 1, \ldots, N-1.$$

Proof. We fix some $n \in \{0, \ldots, N-1\}$. For a each $k \in \{0, \ldots, N-1\}$, we have the discrete Fourier transform given by

$$\widehat{f}_k = \sum_{m=0}^{N-1} f_m \exp\left(-i2\pi m \tfrac{k}{N}\right),$$

where $\exp z = e^z$ is the complex exponential function. Solving for f_n in each of these, we obtain the system of N equations

$$f_n = \widehat{f}_k \exp\left(i2\pi n \tfrac{k}{N}\right) - \sum_{\substack{m=0,\\ m\neq n}}^{N-1} f_m \exp\left(i2\pi(n-m)\tfrac{k}{N}\right), \quad k = 0, 1, \ldots, N-1.$$

Combining these, we obtain

$$N f_n = \sum_{k=0}^{N-1} \widehat{f}_k \exp\left(i2\pi n \tfrac{k}{N}\right) - \sum_{k=0}^{N-1} \sum_{\substack{m=0,\\ m\neq n}}^{N-1} f_m \exp\left(i2\pi(n-m)\tfrac{k}{N}\right). \tag{1}$$

Now, inspecting the summation on the right, we have

$$\sum_{k=0}^{N-1} \sum_{\substack{m=0,\\ m\neq n}}^{N-1} f_m \exp\left(i2\pi(n-m)\tfrac{k}{N}\right) = \sum_{\substack{m=0,\\ m\neq n}}^{N-1} f_m \sum_{k=0}^{N-1} \exp\left(i2\pi(n-m)\tfrac{k}{N}\right).$$

Now since $e^{zk} = (e^z)^k$ for $k \in \mathbb{Z}$, then we have the partial sum of a geometric series

$$\sum_{k=0}^{N-1} \exp\left(i2\pi(n-m)\tfrac{k}{N}\right) = \sum_{k=0}^{N-1} \exp\left(i2\pi(n-m)\tfrac{1}{N}\right)^k,$$

whose sum is given by

$$\frac{1-\exp\left(i2\pi(n-m)\frac{1}{N}\right)^N}{1-\exp\left(i2\pi(n-m)\frac{1}{N}\right)} = \frac{1-\exp(i2\pi(n-m))}{1-\exp\left(i2\pi(n-m)\frac{1}{N}\right)} = \frac{1-1}{1-\exp\left(i2\pi(n-m)\frac{1}{N}\right)} = 0.$$

Note that the second equality above follows from the fact that $n-m$ is an integer, and the exponential of any integer multiple of $i2\pi$ is 1. Substituting back into (1) and dividing by N we obtain the desired equation

$$f_n = \frac{1}{N}\sum_{k=0}^{N-1} \widehat{f}_k \exp\left(i2\pi n\tfrac{k}{N}\right).$$

□

3.2

Show that

$$\widehat{(f*g)}_k = \widehat{f}_k\widehat{g}_k.$$

Proof. Applying first the definition of the discrete Fourier transform, then the result of 3.1, we obtain

$$\widehat{(f*g)}_k = \sum_{m=0}^{N-1}(f*g)_m \exp\left(-i2\pi m\tfrac{k}{N}\right) = \sum_{m=0}^{N-1}\sum_{s=0}^{N-1} f_{m-s}g_s \exp\left(-i2\pi m\tfrac{k}{N}\right).$$

We rearrange the summation and apply the definition of the discrete Fourier transform,

$$\begin{aligned}
\widehat{(f*g)}_k &= \sum_{s=0}^{N-1} g_s \sum_{m=0}^{N-1} f_{m-s}\exp\left(-i2\pi m\tfrac{k}{N}\right) \\
&= \sum_{s=0}^{N-1} g_s \exp\left(-i2\pi s\tfrac{k}{N}\right) \sum_{m=0}^{N-1} f_{m-s}\exp\left(-i2\pi (m-s)\tfrac{k}{N}\right) \\
&= \sum_{s=0}^{N-1} g_s \exp\left(-i2\pi s\tfrac{k}{N}\right) \widehat{f}_k \\
&= \widehat{f}_k \sum_{s=0}^{N-1} g_s \exp\left(-i2\pi s\tfrac{k}{N}\right) \\
&= \widehat{f}_k\widehat{g}_k.
\end{aligned}$$

□

4

The **normalized** Chebyshev polynomials on $x \in [-1,1]$ are defined as $T_n(x) = \sqrt{2/\pi}\cos(n\cos^{-1}x)$, for $n \geq 1$ and $T_0(x) = \sqrt{1/\pi}$. Consider the expansion

$$S_n(x) = \sum_{k=0}^{n} a_k T_k(x),$$

where $\{T_0, T_1, \dots, T_n\}$ is a finite set of Chebyshev polynomials on $[-1,1]$.

4.1

Show the **normalized** Chebyshev polynomials are orthonormal with $\langle T_m, T_n\rangle_\omega = \delta_{m,n}$ under the inner-product

$$\langle u, v\rangle_\omega = \int_{-1}^{1} u(x)v(x)\omega(x)\,dx,$$

where $\omega(x) = \frac{1}{\sqrt{1-x^2}}$.

Proof. For any pair m, n, we have

$$\langle T_m, T_n\rangle_\omega = \int_{-1}^{1} T_m(x)T_n(x)\frac{dx}{\sqrt{1-x^2}}.$$

On this interval, $\cos^{-1}x$ is a well-defined inverse of $\cos x$. Taking $\theta = \cos^{-1}x$, we rewrite the integration as

$$\langle T_m, T_n\rangle_\omega = \int_{\cos\pi}^{\cos 0} T_m(\cos\cos^{-1}x)T_n(\cos\cos^{-1}x)\frac{dx}{\sqrt{1-x^2}} = \int_0^\pi T_m(\cos\theta)T_n(\cos\theta)\,d\theta.$$

In particular, we have

$$\langle T_0, T_0\rangle_\omega = \int_0^\pi \frac{1}{\pi}\,d\theta = 1.$$

And for any $m > 0$, we find

$$\langle T_m, T_0\rangle_\omega = \frac{\sqrt{2}}{\pi}\int_0^\pi \cos(m\theta)\,d\theta = \frac{\sqrt{2}}{\pi}\left(\frac{\sin(m\pi)}{m} - \frac{\sin(m0)}{m}\right) = 0.$$

Now, suppose $m, n \geq 1$, then

$$\begin{aligned}\langle T_m, T_n\rangle_\omega &= \frac{2}{\pi}\int_0^\pi \cos(m\theta)\cos(n\theta)\,d\theta \\ &= \frac{1}{\pi}\left(\int_0^\pi \cos((m+n)\theta)\,d\theta + \int_0^\pi \cos((m-n)\theta)\,d\theta\right).\end{aligned}$$

Regardless of whether m and n are equal, we have

$$\int_0^\pi \cos((m+n)\theta)\,d\theta = \frac{\sin((m+n)\pi)}{m+n} - \frac{\sin((m+n)0)}{m+n} = 0.$$

Then if $m = n$, then

$$\langle T_m, T_n \rangle_\omega = \frac{1}{\pi} \int_0^\pi \cos((m-n)\theta)\, d\theta = \frac{1}{\pi} \int_0^\pi \cos 0\, d\theta = 1.$$

And if $m \neq n$, then

$$\langle T_m, T_n \rangle_\omega = \frac{1}{\pi} \int_0^\pi \cos((m-n)\theta)\, d\theta = \frac{1}{\pi} \left(\frac{\sin((m-n)\pi)}{m-n} - \frac{\sin((m-n)0)}{m-n} \right) = 0.$$

Thus, for all $m, n \geq 0$, we have $\langle T_m, T_n \rangle_\omega = \delta_{m,n}$.

□

4.2

To approximate the function $f(x) \in C[-1, 1]$, find the coefficients a_k, $k = 0, \ldots, n$ so that $S_n(x)$ minimizes the least square error

$$E(a_0, \ldots, a_n) = \int_{-1}^1 (f(x) - S_n(x))^2 \omega(x)\, dx,$$

where $\omega(x) \frac{1}{\sqrt{1-x^2}}$. You are required to show the derivation process.

Note that the set of Chebyshev polynomial $\{T_0, \ldots, T_n\}$ form a basis for the space P_n of real polynomials with degree at most n. This is because the dimension of P_n is $n+1$, and the set of $n+1$ Chebyshev polynomials are orthonormal (In particular, they form an orthonormal basis). We denote by $\|\cdot\|_\omega$ the norm induced by the weighted inner product,

$$\|u\|_\omega = \sqrt{\langle u, u \rangle_\omega}.$$

Then if $(a_0, \ldots, a_n)$ are the coefficients of S_n as a combination of Chebyshev polynomials, the least square error is given by

$$E(a_0, \ldots, a_n) = \|f - S_n\|_\omega^2.$$

Proposition 1. *If $\langle f - S_n, p \rangle_\omega = 0$ for all $p \in P_n$, then S_n minimizes $\|f - S_n\|_\omega^2$ in P_n.*

Proof. Suppose $S_n \in P_n$ such that $f - S_n$ is orthogonal to P_n. Let $p \in P_n$, then we have

$$\|f - p\|_\omega^2 = \|f - S_n + S_n - p\|_\omega^2.$$

Because $f - S_n$ is orthogonal to P_n, in particular it is orthogonal to $S_n - p \in P_n$. Then

$$\|f - p\|_\omega^2 = \|f - S_n\|_\omega^2 + \|S_n - p\|_\omega^2 \geq \|f - S_n\|_\omega^2.$$

Thus, S_n minimizes the error in P_n.

□

Because the set of Chebyshev polynomials forms a basis for P_n, then it is sufficient to find S_n such that $f - S_n$ is orthogonal to each Chebyshev polynomial. That is, for each $k \in \{0, \dots, n\}$, we want

$$0 = \langle f - S_n, T_k \rangle_\omega = \langle f, T_k \rangle_\omega - \langle S_n, T_k \rangle_\omega .$$

Expanding the rightmost product, we find

$$\langle S_n, T_k \rangle_\omega = \langle a_0 T_0 + \cdots + a_n T_n, T_k \rangle_\omega = \sum_{\ell=0}^{n} a_\ell \langle T_\ell, T_k \rangle = \sum_{\ell=0}^{n} a_\ell \delta_{\ell,k} = a_k.$$

Thus, the coefficients

$$a_k = \langle f, T_k \rangle_\omega = \int_{-1}^{1} f(x) T_k(x) \omega(x)\, dx, \quad k = 0, \dots, n,$$

minimize the least square error $E(a_0, \dots, a_n)$.

This shows that, in general, given an inner product space of functions, the best approximation of that function in a particular subspace, with respect to the induced norm, will be the projection of that function onto the subspace. In this case, finding the projection of f onto P_n is helped by having the orthonormal basis for P_n of Chebyshev polynomials.