1. Solve the steady state heat conduction problem from Question 4 from HW#4 using the conjugate gradient method. Take the same $h = 2^{-p}$, $p = 2, 3, 4, 5$ and start from $\mathbf{x}^{(0)} = 0$. Terminate the iteration when $\|r^{(k)}\|_2 \le 10^{-2}$, where $r^{(k)}$ is the residual at step k. Display the results for each method in a table:

   - column 1: h (mesh size)
   - column 2: k (total number of iterations needed to achieve error tolerance)
   - column 3: $\|r^{(k)}\|_2$ (norm of residual at step k)
   - column 4: $\dfrac{\|r^{(k)}\|_2}{\|r^{(k-1)}\|_2}$ (ratio of successive residual norms at step k)

   Summarize and explain the results. How does the rate of convergence behave as h is reduced? How do the results compare with those obtained in Question 4 from HW#4?

   From the last homework, we have the following tables for Jacobi and Gauss-Seidel method:

   Table 1: Jacobi results

h	k (iters)	$\|r_k\|_2$	ratio
0.2500	13	9.568319×10^{-3}	7.071067×10^{-1}
0.1250	49	9.945191×10^{-3}	9.238773×10^{-1}
0.0625	166	9.887564×10^{-3}	9.807788×10^{-1}
0.0312	523	9.998456×10^{-3}	9.951703×10^{-1}

   Table 2: Gauss–Seidel results

h	k (iters)	$\|r_k\|_2$	ratio
0.2500	7	7.768911×10^{-3}	4.999220×10^{-1}
0.1250	24	9.751416×10^{-3}	8.532369×10^{-1}
0.0625	81	9.695216×10^{-3}	9.617785×10^{-1}
0.0312	256	9.937994×10^{-3}	9.903053×10^{-1}

   For the Conjugate Gradient method, the results are as follows:

   Table 3: Conjugate Gradient results

h	k (iters)	$\|r_k\|_2$	ratio
0.2500	5	1.726279×10^{-17}	0
0.1250	11	8.756590×10^{-3}	2.536811×10^{-1}
0.0625	23	6.355795×10^{-3}	3.777855×10^{-1}
0.0312	46	7.127013×10^{-3}	5.431558×10^{-1}

   Summary:

   Behavior of CG method as h decreases:

   - Iterations roughly double as h halves, so CG iteration count grows like $O(1/h)$.
   - The ratio of successive residual norms increases as h decreases, $0 \to 0.254 \to 0.378 \to 0.543$, indicating slower convergence. This is consistent with theory:

   $$\rho_{CG} = \frac{\sqrt{\kappa_2} - 1}{\sqrt{\kappa_2} + 1} \to 1 \text{ as } \kappa \to \infty,$$

   Comparison with Jacobi and Gauss-Seidel:

   - CG needs far fewer iterations than Jacobi and Gauss-Seidel for the same h. For example, at $h = 0.0312$, CG takes 46 iterations vs. 523 (Jacobi) and 256 (Gauss-Seidel).
   - For residual contraction, CG shows slower approaching to 1 as h decreases, compared with Jacobi and Gauss-Seidel.

2. Consider the method of steepest descent for solving $Ax = b$, where A is symmetric, positive definite.

(a) Derive the relation
$$r^{(k)} = r^{(k-1)} - \alpha^{(k)} A r^{(k-1)}.$$

(b) Show that consecutive residuals are orthogonal, i.e.,
$$\left(r^{(k)}\right)^\top r^{(k-1)} = 0.$$

Solution:

(a) By definition, the residual at step k is $r^{(k)} = b - Ax^{(k)}$. The update step in the method of steepest descent is given by
$$x^{(k)} = x^{(k-1)} + \alpha^{(k)} r^{(k-1)}.$$
Multiplying by A and subtracting from b, we get
$$\begin{aligned} b - Ax^{(k)} &= b - A\left(x^{(k-1)} + \alpha^{(k)} r^{(k-1)}\right) \\ r^{(k)} &= (b - Ax^{(k-1)}) - \alpha^{(k)} A r^{(k-1)} \\ r^{(k)} &= r^{(k-1)} - \alpha^{(k)} A r^{(k-1)}. \end{aligned}$$

(b) We compute the inner product $(r^{(k)})^\top r^{(k-1)}$ using the relation derived in (a):
$$(r^{(k)})^\top r^{(k-1)} = \left(r^{(k-1)} - \alpha^{(k)} A r^{(k-1)}\right)^\top r^{(k-1)}.$$
Distributing the transpose and the product:
$$(r^{(k)})^\top r^{(k-1)} = (r^{(k-1)})^\top r^{(k-1)} - \alpha^{(k)} (A r^{(k-1)})^\top r^{(k-1)}.$$
Since A is symmetric, $(Ar^{(k-1)})^\top = (r^{(k-1)})^\top A^\top = (r^{(k-1)})^\top A$. Thus,
$$(r^{(k)})^\top r^{(k-1)} = (r^{(k-1)})^\top r^{(k-1)} - \alpha^{(k)} (r^{(k-1)})^\top A r^{(k-1)}.$$
In the method of steepest descent, the step size $\alpha^{(k)}$ is chosen to minimize the function $\phi(x^{(k-1)} + \alpha r^{(k-1)})$, which yields the formula:
$$\alpha^{(k)} = \frac{(r^{(k-1)})^\top r^{(k-1)}}{(r^{(k-1)})^\top A r^{(k-1)}}.$$
Substituting this into the expression for the inner product:
$$(r^{(k)})^\top r^{(k-1)} = (r^{(k-1)})^\top r^{(k-1)} - \frac{(r^{(k-1)})^\top r^{(k-1)}}{(r^{(k-1)})^\top A r^{(k-1)}} (r^{(k-1)})^\top A r^{(k-1)}.$$
The terms cancel out:
$$(r^{(k)})^\top r^{(k-1)} = (r^{(k-1)})^\top r^{(k-1)} - (r^{(k-1)})^\top r^{(k-1)} = 0.$$
Therefore, consecutive residuals are orthogonal.

3. Let A be symmetric, positive definite and suppose that $\{p_1, \dots, p_k\}$ are conjugate.

(a) Show that $\{p_1, \dots, p_k\}$ are linearly independent.

Proof. Since $\{p_1, \dots, p_k\}$ are A-conjugate, we have
$$p_i^\top A p_j = 0 \quad \text{for } i \neq j.$$
Assume there exist scalars $\alpha_1, \alpha_2, \dots, \alpha_k$ not all zero such that
$$\alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_k p_k = 0.$$
Pick $a_j \neq 0$, and multiply both sides by $p_j^\top A$:
$$p_j^\top A(\alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_k p_k) = 0.$$
Using the conjugacy property, all terms vanish except for the j-th term:
$$\alpha_j p_j^\top A p_j = 0.$$
Since A is positive definite, $p_j^\top A p_j > 0$, which implies $\alpha_j = 0$. This contradicts our assumption that not all α_i are zero. Therefore, the set $\{p_1, \dots, p_k\}$ is linearly independent. □

(b) Show that $\{Ap_1, \dots, Ap_k\}$ are linearly independent.

Proof. Assume there exist scalars $\beta_1, \beta_2, \dots, \beta_k$ not all zero such that

$$\beta_1 Ap_1 + \beta_2 Ap_2 + \cdots + \beta_k Ap_k = 0.$$

Pick $b_j \neq 0$, and multiply both sides by $p_j^\top$:

$$p_j^\top (\beta_1 Ap_1 + \beta_2 Ap_2 + \cdots + \beta_k Ap_k) = 0.$$

Using the conjugacy property, all terms vanish except for the j-th term:

$$\beta_j p_j^\top Ap_j = 0.$$

Since A is positive definite, $p_j^\top Ap_j > 0$, which implies $\beta_j = 0$. This contradicts our assumption that not all β_i are zero. Therefore, the set $\{Ap_1, \dots, Ap_k\}$ is linearly independent. □

(c) Show that $P_k^\top AP_k$ is invertible, where $P_k = [\, p_1 \ \cdots \ p_k \,]$.

Proof. Consider the (i, j) entry of the matrix $P_k^\top AP_k$:

$$(P_k^\top AP_k)_{ij} = p_i^\top Ap_j.$$

Using the conjugacy property, we have

$$(P_k^\top AP_k)_{ij} = \begin{cases} p_i^\top Ap_i > 0 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

So the matrix $P_k^\top AP_k$ is diagonal with positive entries on the diagonal, which implies it is invertible. □

4. Prove that the Lagrange polynomials $\ell_j \in \mathcal{P}_n$ form a basis for $\mathcal{P}_n$.

Proof. Recall that the Lagrange polynomials $\ell_j(x)$ are defined as

$$\ell_j(x) = \prod_{\substack{i=0 \\ i \neq j}}^{n} \frac{x - x_i}{x_j - x_i},$$

where $x_0, x_1, \dots, x_n$ are distinct nodes.

- **Linearly Independent:**
  Assume there exist scalars $\alpha_0, \alpha_1, \dots, \alpha_n$ such that

  $$\alpha_0 \ell_0(x) + \alpha_1 \ell_1(x) + \cdots + \alpha_n \ell_n(x) = 0.$$

  Evaluating this equation at $x = x_j$ for each $j = 0, 1, \dots, n$, we get

  $$\alpha_j = 0.$$

  Since this holds for all j, we conclude that all coefficients must be zero, proving linear independence.
- **Spanning:** Let $p(x) \in \mathcal{P}_n$ be an arbitrary polynomial of degree at most n. We can express $p(x)$ in terms of the Lagrange polynomials as follows:

  $$p(x) = \sum_{j=0}^{n} p(x_j) \ell_j(x).$$

  This shows that any polynomial in $\mathcal{P}_n$ can be expressed as a linear combination of the Lagrange polynomials, proving that they span $\mathcal{P}_n$.

□

5. Prove that for the Lagrange polynomials $\ell_j \in \mathcal{P}_n$,

$$\sum_{j=0}^{n} \ell_j(x) \equiv 1.$$

Proof.

$$\ell_j(x) = \prod_{\substack{i=0 \\ i \neq j}}^{n} \frac{x - x_i}{x_j - x_i},$$

where $x_0, x_1, \ldots, x_n$ are distinct nodes. Consider the sum

$$S(x) = \sum_{j=0}^{n} \ell_j(x).$$

We will show that $S(x) \equiv 1$. First, note that each $\ell_j(x)$ is a polynomial of degree n. Therefore, $S(x)$ is also a polynomial of degree at most n. Next, evaluate $S(x)$ at each of the nodes x_k for $k = 0, 1, \ldots, n$:

$$S(x_k) = \sum_{j=0}^{n} \ell_j(x_k).$$

By the definition of the Lagrange polynomials, we have

$$\ell_j(x_k) = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

Therefore,

$$S(x_k) = \ell_k(x_k) + \sum_{\substack{j=0 \\ j \neq k}}^{n} \ell_j(x_k) = 1 + 0 = 1.$$

Since $S(x)$ is a polynomial of degree at most n and takes the value 1 at $n+1$ distinct points $x_0, x_1, \ldots, x_n$, it must be the constant polynomial 1. Thus,

$$S(x) \equiv 1.$$

□

6. Prove that

$$\omega'_{n+1}(x_i) = \prod_{\substack{j=0 \\ j \neq i}}^{n} (x_i - x_j),$$

where ω_{n+1} is the nodal polynomial.

Proof. The nodal polynomial is defined as $\omega_{n+1}(x) = \prod_{j=0}^{n}(x - x_j)$. We can rewrite this as

$$\omega_{n+1}(x) = (x - x_i) \prod_{\substack{j=0 \\ j \neq i}}^{n} (x - x_j) = (x - x_i) g(x),$$

where $g(x) = \prod_{j \neq i}(x - x_j)$. Differentiating using the product rule gives

$$\omega'_{n+1}(x) = 1 \cdot g(x) + (x - x_i) g'(x).$$

Evaluating at $x = x_i$, the second term vanishes:

$$\omega'_{n+1}(x_i) = g(x_i) + 0 = \prod_{\substack{j=0 \\ j \neq i}}^{n} (x_i - x_j).$$

□

7. Provide an estimate of $\|\omega_{n+1}\|_\infty$ in the cases $n = 1$ and $n = 2$, for a distribution of equally spaced nodes.

Proof. For $n = 1$, the nodal polynomial is

$$\omega_2(x) = (x - x_0)(x - x_1).$$

Assuming equally spaced nodes $x_0 = a$ and $x_1 = b$, we have

$$\omega_2(x) = (x - a)(x - b).$$

The maximum value of $|\omega_2(x)|$ occurs at the midpoint $x = \frac{a+b}{2}$:

$$\|\omega_2\|_\infty = \left|\left(\frac{a+b}{2} - a\right)\left(\frac{a+b}{2} - b\right)\right| = \left|\frac{b-a}{2} \cdot \left(-\frac{b-a}{2}\right)\right| = \frac{(b-a)^2}{4}.$$

For $n = 2$, the nodal polynomial is

$$\omega_3(x) = (x - x_0)(x - x_1)(x - x_2).$$

Assuming equally spaced nodes $x_0 = a$, $x_1 = a + h$, and $x_2 = a + 2h$, we have

$$\omega_3(x) = (x - a)(x - (a + h))(x - (a + 2h)).$$

$$\begin{aligned}
\omega_3'(x) &= (x - (a + h))(x - (a + 2h)) + (x - a)(x - (a + 2h)) + (x - a)(x - (a + h)) \\
&= 3x^2 - 6(a + h)x + (3a^2 + 6ah + 2h^2) \\
&= 3\left(x^2 - 2(a + h)x + a^2 + 2ah + h^2\right) - h^2 \\
&= 3\left(x - (a + h)\right)^2 - h^2 = 0 \implies x = a + h \pm \frac{h}{\sqrt{3}}.
\end{aligned}$$

Evaluating $\omega_3(x)$ at these critical points:

$$\begin{aligned}
\omega_3\left(a + h + \frac{h}{\sqrt{3}}\right) &= \left(h + \frac{h}{\sqrt{3}}\right)\left(\frac{h}{\sqrt{3}}\right)\left(-h + \frac{h}{\sqrt{3}}\right) = -\frac{2h^3}{3\sqrt{3}}, \\
\omega_3\left(a + h - \frac{h}{\sqrt{3}}\right) &= \left(h - \frac{h}{\sqrt{3}}\right)\left(-\frac{h}{\sqrt{3}}\right)\left(-h - \frac{h}{\sqrt{3}}\right) = \frac{2h^3}{3\sqrt{3}}.
\end{aligned}$$

Thus,

$$\|\omega_3\|_\infty = \frac{2h^3}{3\sqrt{3}} = \frac{2\sqrt{3}h^3}{9} = \frac{\sqrt{3}}{36}(b - a)^3.$$

□

8. Some modeling considerations have mandated a search for a function

$$u(x) = \gamma_0 e^{\gamma_1 x + \gamma_2 x^2},$$

where the unknown coefficients γ_1 and γ_2 are expected to be nonpositive. Given are data pairs to be interpolated, (x_0, z_0), (x_1, z_1), and (x_2, z_2), where $z_i > 0$, $i = 0, 1, 2$. Thus, we require

$$u(x_i) = z_i.$$

The function $u(x)$ is not linear in its coefficients, but

$$v(x) = \ln(u(x))$$

is linear in its coefficients.

(a) Find a quadratic polynomial $v(x)$ that interpolates appropriately defined three data pairs, and then give a formula for $u(x)$ in terms of the original data. (This is a pen-and-paper item; the following one should consume much less of your time.)

Solution: We have

$$v(x) = \ln(u(x)) = \ln(\gamma_0) + \gamma_1 x + \gamma_2 x^2.$$

And the data pairs for $v(x)$ are $(x_i, \ln(z_i))$, $i = 0, 1, 2$. We can express $v(x)$ in terms of Lagrange basis polynomials:

$$v(x) = \sum_{j=0}^{2} \ln(z_j)\ell_j(x),$$

where $\ell_j(x)$ are the Lagrange basis polynomials defined as

$$\ell_j(x) = \prod_{\substack{i=0 \\ i\neq j}}^{2} \frac{x - x_i}{x_j - x_i}.$$

Therefore, the function $u(x)$ can be expressed as

$$u(x) = \exp\left(\sum_{j=0}^{2} \ln(z_j)\ell_j(x)\right) = \prod_{j=0}^{2} z_j^{\ell_j(x)}.$$

(b) Write a script to find u for the data $(0, 1)$, $(1, 0.9)$, $(3, 0.5)$. Give the coefficients γ_i and plot the resulting interpolant over the interval $[0, 6]$. In what way does the curve behave qualitatively differently from a quadratic?

Using a MATLAB script, we find the coefficients:

$$\gamma_0 = 1$$
$$\gamma_1 = -0.042516$$
$$\gamma_2 = -0.062844$$

The plot of the interpolant $u(x)$ over the interval $[0, 6]$ shows an exponential decay behavior, which is qualitatively different from a quadratic function that would typically exhibit a parabolic shape. The exponential form allows for a more gradual decrease, especially as x increases, compared to the rapid changes seen in quadratic functions.

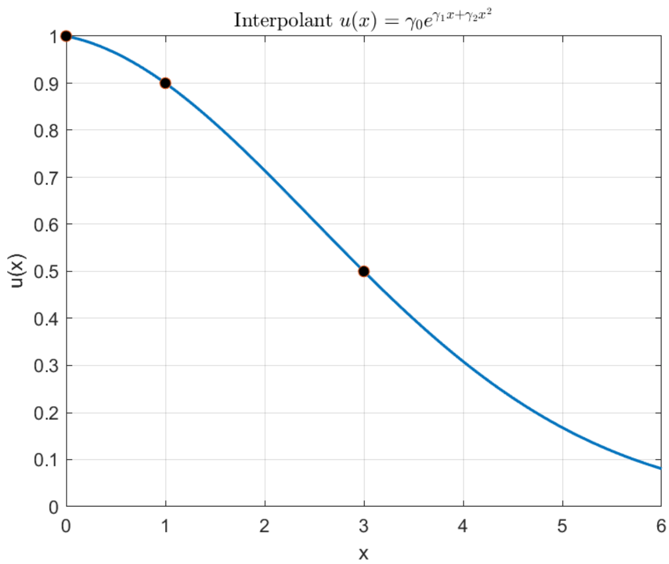

Figure 1: Plot of the interpolant $u(x)$ over $[0, 6]$

9. Use the known values of the function $\sin(x)$ at

$$x = 0,\ \frac{\pi}{6},\ \frac{\pi}{4},\ \frac{\pi}{3},\ \frac{\pi}{2}$$

to derive an interpolating polynomial $p(x)$. What is the degree of your polynomial? What is the interpolation error magnitude $|p(1.2) - \sin(1.2)|$?

Solution: The degree of the interpolating polynomial is 4, since we have 5 data points. Using Lagrange interpolation, we construct the polynomial $p(x)$ as follows:

$$p(x) = \sum_{j=0}^{4} \sin(x_j)\ell_j(x),$$

where $x_0 = 0$, $x_1 = \frac{\pi}{6}$, $x_2 = \frac{\pi}{4}$, $x_3 = \frac{\pi}{3}$, $x_4 = \frac{\pi}{2}$, and $\ell_j(x)$ are the Lagrange basis polynomials defined as

$$\ell_j(x) = \prod_{\substack{i=0 \\ i\neq j}}^{4} \frac{x - x_i}{x_j - x_i}.$$

Evaluating $p(1.2)$ and comparing it to $\sin(1.2)$:

$$\begin{gathered} p(1.2) = 0.9321 \\ \sin(1.2) = 0.932039 \\ |p(1.2) - \sin(1.2)| = 1.0244 \times 10^{-4} \end{gathered}$$

10. Let $x_0, \ldots, x_n$ be distinct real points, and consider the following interpolation problem. Choose a function

$$P_n(x) = \sum_{j=0}^{n} c_j e^{jx}$$

such that

$$P_n(x_i) = y_i, \quad i = 0, 1, \ldots, n,$$

with the $\{y_i\}$ given data. Show there is a unique choice of $c_0, \ldots, c_n$.

Proof. Suppose there are two different sets of coefficients $\{c_j\}$ and $\{d_j\}$ that satisfy the interpolation conditions:

$$P_n(x_i) = \sum_{j=0}^{n} c_j e^{jx_i} = \sum_{j=0}^{n} d_j e^{jx_i} = y_i, \quad i = 0, 1, \ldots, n,$$

Subtracting these two equations, we get

$$\sum_{j=0}^{n} (c_j - d_j)e^{jx_i} = 0, \quad i = 0, 1, \ldots, n.$$

Define polynomial

$$g(z) := \sum_{j=0}^{n} a_j z^j,$$

Then take $z = e^x$,

$$g(z_i) = 0 \text{ where } z_i = e^{x_i}, i = 0, 1, \ldots, n.$$

But g is a polynomial of degree at most n with $n + 1$ distinct roots $z_0, z_1, \ldots, z_n$. By the Fundamental Theorem of Algebra, the only possible case is that

$$g(z) \equiv 0.$$

Hence all coefficients $a_j = c_j - d_j = 0$, which implies $c_j = d_j$ for all $j = 0, 1, \ldots, n$. □

11. Consider finding a rational function

$$p(x) = \frac{a + bx}{1 + cx}$$

that satisfies

$$p(x_i) = y_i, \quad i = 1, 2, 3,$$

with x_1, x_2, x_3 distinct. Does such a function $p(x)$ exist, or are additional conditions needed to ensure existence and uniqueness of $p(x)$?

Solution: The interpolation condition $p(x_i) = y_i$ implies

$$y_i = \frac{a + bx_i}{1 + cx_i} \implies y_i(1 + cx_i) = a + bx_i \implies a + bx_i - cx_i y_i = y_i.$$

For $i = 1, 2, 3$, this gives a linear system of equations for the unknowns a, b, c:

$$\begin{pmatrix} 1 & x_1 & -x_1 y_1 \\ 1 & x_2 & -x_2 y_2 \\ 1 & x_3 & -x_3 y_3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

Let M be the coefficient matrix. A unique solution for a, b, c exists if and only if $\det(M) \neq 0$.

However, We must also ensure that the denominator does not vanish at any of the interpolation nodes x_i. That is, we require

$$1 + cx_i \neq 0 \quad \text{for } i = 1, 2, 3.$$

Therefore, such a function $p(x)$ does not always exist. Additional conditions are needed:

(a) The linear system must be invertible ($\det(M) \neq 0$).

(b) The resulting c must satisfy $1 + cx_i \neq 0$ for all $i = 1, 2, 3$.

12. The following data are taken from a polynomial of degree ≤ 5:

$$(-2, -5),\ (-1, 1),\ (0, 1),\ (1, 1),\ (2, 7),\ (3, 25).$$

What is the degree of the polynomial?

Solution: Consider a difference table for the given data points. Since the x values are equally spaced with $h = 1$, we can use forward differences.

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	$\Delta^5 y$
-2	-5					
		6				
-1	1		-6			
		0		6		
0	1		0		0	
		0		6		0
1	1		6		0	
		6		6		
2	7		12			
		18				
3	25					

The third differences ($\Delta^3 y$) are constant (all equal to 6). This implies that the polynomial is of degree **3**.

13. For $f(x) = \dfrac{1}{1 + x^2}$, $-5 \leq x \leq 5$, produce $p_n(x)$ using $n + 1$ evenly spaced nodes on $[-5, 5]$. Calculate $p_n(x)$ at a large number of points, and graph it and its error on $[-5, 5]$.

We choose $n = 10$ for the interpolation. The result shows significant oscillations near the endpoints of the interval, a phenomenon known as Runge's phenomenon. The error plot indicates that the maximum error occurs near the edges of the interval.

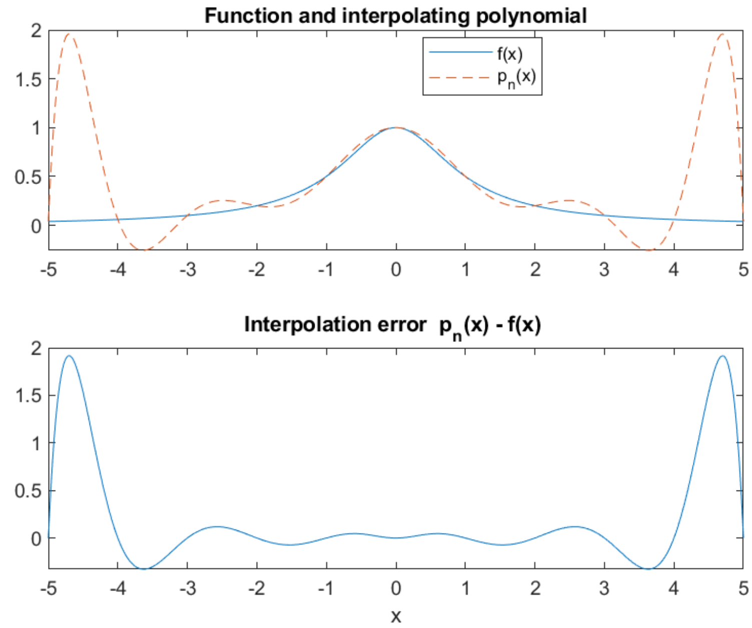

Figure 2: Interpolation of $f(x) = \frac{1}{1+x^2}$ using $n+1 = 11$ evenly spaced nodes on $[-5, 5]$

14. Given a sequence $y_0, y_1, y_2, \dots$, define the forward difference operator Δ by

$$\Delta y_i := y_{i+1} - y_i.$$

Powers of Δ are defined recursively by

$$\Delta^0 y_i := y_i, \qquad \Delta^j y_i := \Delta(\Delta^{j-1} y_i), \quad j = 1, 2, \dots.$$

Thus,

$$\Delta^2 y_i = \Delta(y_{i+1} - y_i) = y_{i+2} - 2y_{i+1} + y_i,$$

etc.

Consider the polynomial interpolation at equispaced points

$$x_i = x_0 + ih, \quad i = 0, 1, \dots, n.$$

(a) Show that

$$f[x_0, x_1, \dots, x_j] = \frac{1}{j!\, h^j} \Delta^j f(x_0).$$

Hint: Use mathematical induction.

Proof. We will prove the statement by induction on j.

- For j=0:

$$f[x_0] = f(x_0) = \Delta^0 f(x_0)$$

- For j=1:

$$f[x_0, x_1] = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = \frac{f(x_0 + h) - f(x_0)}{h} = \frac{\Delta f(x_0)}{1!\, h^1}$$

- Assume the statement holds for j points:

$$f[x_0, x_1, \dots, x_{j-1}] = \frac{1}{(j-1)!\, h^{j-1}} \Delta^{j-1} f(x_0).$$

$$f[x_1, x_2, \dots, x_j] = \frac{1}{(j-1)!\, h^{j-1}} \Delta^{j-1} f(x_1).$$

So,

$$\begin{aligned} f[x_0, x_1, \ldots, x_j] &= \frac{f[x_1, x_2, \ldots, x_j] - f[x_0, x_1, \ldots, x_{j-1}]}{x_j - x_0} \\ &= \frac{\frac{1}{(j-1)!\, h^{j-1}}\, \Delta^{j-1} f(x_1) - \frac{1}{(j-1)!\, h^{j-1}}\, \Delta^{j-1} f(x_0)}{jh} \\ &= \frac{1}{j!\, h^j} \left(\Delta^{j-1} f(x_1) - \Delta^{j-1} f(x_0) \right) \\ &= \frac{1}{j!\, h^j} \Delta^j f(x_0). \end{aligned}$$

□

(b) Show that the interpolating polynomial of degree at most n is given by the Newton forward difference formula

$$p_n(x) = \sum_{j=0}^{n} \binom{s}{j} \Delta^j f(x_0),$$

where

$$s = \frac{x - x_0}{h}$$

and

$$\binom{s}{j} = \frac{s(s-1)\cdots(s-j+1)}{j!}, \qquad \binom{s}{0} = 1.$$

Proof. The Newton form of the interpolating polynomial is given by

$$p_n(x) = \sum_{j=0}^{n} f[x_0, x_1, \ldots, x_j] \prod_{i=0}^{j-1} (x - x_i).$$

For equal spacing, we have

$$x - x_k = (x - x_0) - kh = h\left(\frac{x - x_0}{h} - k\right)$$

Define $s = \frac{x - x_0}{h}$, then

$$\prod_{i=0}^{j-1} (x - x_i) = h^j \prod_{i=0}^{j-1} (s - i) = h^j \frac{s(s-1)\cdots(s-j+1)}{j!} j! = h^j \binom{s}{j} j!.$$

Using the result from part (a), we substitute $f[x_0, x_1, \ldots, x_j] = \frac{1}{j!\, h^j} \Delta^j f(x_0)$ into the Newton form:

$$\begin{aligned} p_n(x) &= \sum_{j=0}^{n} \left(\frac{1}{j!\, h^j} \Delta^j f(x_0) \right) \left(h^j \binom{s}{j} j! \right) \\ &= \sum_{j=0}^{n} \binom{s}{j} \Delta^j f(x_0). \end{aligned}$$

□