1 Polynomial Interpolation

Given \( n \) distinct points, we can find a unique degree \( n-1 \) polynomial that passes through these points. Let the polynomial \( p \) be

\[
p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-1} x^{n-1}.
\]

Let the \( n \) points be

\[
p(x_1) = y_1, p(x_2) = y_2, \ldots, p(x_n) = y_n,
\]

where \( x_1 \neq x_2 \neq \cdots \neq x_n \).

We can construct a matrix-vector equation as follows to recover the polynomial \( p \).

\[
\begin{bmatrix}
1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\
1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & x_n & x_n^2 & \cdots & x_n^{n-1}
\end{bmatrix}
\begin{bmatrix}
a_0 \\
a_1 \\
\vdots \\
a_{n-1}
\end{bmatrix}
= 
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n
\end{bmatrix}
\]

We can solve for the \( a \) values by setting:

\[
\bar{a} = A^{-1} \bar{y}
\]

Note that the matrix \( A \) is known as a Vandermonde matrix whose determinant is given by

\[
\text{det}(A) = \prod_{1 \leq i < j \leq n} (x_j - x_i)
\]

Since \( x_1 \neq x_2 \neq \cdots \neq x_n \), the determinant is non-zero and \( A \) is always invertible.

2 Polynomial Regression

Sometimes we may want to fit our data to a polynomial with an order less than \( n-1 \). If we fit the data to a polynomial of order \( m < n \) we get:

\[
p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{m-1} x^{m-1}
\]

Now when we construct the matrix-vector equation to recover polynomial \( p \), we get:

\[
\begin{bmatrix}
1 & x_1 & x_1^2 & \cdots & x_1^{m-1} \\
1 & x_2 & x_2^2 & \cdots & x_2^{m-1} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & x_n & x_n^2 & \cdots & x_n^{m-1}
\end{bmatrix}
\begin{bmatrix}
a_0 \\
a_1 \\
\vdots \\
a_{m-1}
\end{bmatrix}
= 
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_m
\end{bmatrix}
\]

With this matrix equation, we have \( n \) equations with \( m \) unknowns, which means our system is over-defined (since \( m < n \)). One way to find the best fitting \( a \) values for this polynomial is to use least-squares, where you set:

\[
\bar{a} = (A^T A)^{-1} A^T \bar{y}
\]
3 Lagrange Interpolation

In practice, to approximate some unknown or complex function \( f(x) \), we take \( n \) evaluations/samples of the function, denoted by \( \{(x_i, y_i) \in f(x_i)\): \( 0 \leq i \leq n - 1 \}. \) For the rest of this question, we will consider the following three points: \( \{(0, 3), (1, 4), (3, -6)\}. \)

a) Using the interpolation method discussed above, find the matrix \( A \) such that \( A\tilde{a} = \tilde{y}. \)

b) What are the coefficients \( a_0, a_1, a_2? \)

c) Observe that this system very quickly becomes frustrating to solve—as \( n \) increases, the difficulty of calculating the inverse increases far more quickly.

This is where Lagrange interpolation can be useful; the idea of Lagrange interpolation is that, instead of writing the polynomial in question in terms of \( \{1, x, x^2\}, \) we will write it in terms of \( \{L_0(x), L_1(x), L_2(x)\}, \) where each

\[
L_i(x_j) = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases}
\]

With that, the problem reduces to finding these new coefficients \( b_0, b_1, b_2 \) of the function

\[
f(x) = b_0L_0(x) + b_1L_1(x) + b_2L_2(x)
\]

such that \( f(x_i) = y_i, \forall i = 0, 1, 2. \) What are these coefficients \( b_i? \)

d) Show that if we define

\[
L_i(x) = \prod_{j=0; j \neq i}^{n-1} \frac{x - x_j}{x_i - x_j}
\]

then the condition requested from part (c) is satisfied.

e) Based on the previous two parts, write down the explicit form of \( f(x) \) that passes through the samples \( \{(0, 3), (1, 4), (3, -6)\} \) in terms of \( x \) as opposed to \( L_i(x) \). The resulting formula is the so-called Lagrange polynomial which passes through the \( n \) sampled points. Does this agree with the previous method?

f) Now, suppose instead we wanted to use regression to fit our 3 points to a linear system \( f(x) = a_0 + a_1x. \) What are the best-fit coefficients \( a_0 \) and \( a_1 \) in this situation?