Introduction to Numerical Analysis I
Handout 6

1 Interpolation

We represent functions by finite table $(x_i, f(x_i))$, which is not unique and consist irreversible lost of data.

Definition 1.1. Let f be known at points $\{x_i\}_{i=0}^n$. Obtaining the value of f at $x \in [x_0, x_n]$ from the known data $(x_i, f(x_i))$ is called **interpolation**. A function used to generate interpolation is called interpolant.

Obtaining the value of f at $x \notin [x_0, x_n]$ from the known data $(x_i, f(x_i))$ is called **extrapolation**.
The extrapolation is less accurate then interpolation, but other than that it uses the same formulas as interpolation.

We use polynomials for interpolation.

Theorem 1.2. Weierstrass Approximation Theorem Let $f(x)$ be continues function on an interval $[a, b]$ and let $\varepsilon > 0$, then there exists a polynomial $P(x)$ such that $|f(x) - P(x)| < \varepsilon$. Note that it is not required that f be differentiable.

1.1 Polynomials

One of the standard notations for n'th order polynomials is $P_n(x) = \sum\limits_{k=0}^{n} a_k x^k$ or $P_n(x) = \sum\limits_{k=0}^{n} a_k (x-x_0)^k$, where $a_n \neq 0$, and the most general form is

$$P_n(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + ++ \\ a_n(x - x_0) \cdots (x - x_{n-1})$$

Horner's Rule is the efficient method to compute polynomials, given here for the most general third form (it is easy to reduce it for the other forms)

$$(((a_n(x - x_{n-1}) + a_{n-1})(x - x_{n-2}) \\ +a_{n-2})(x - x_{n-3}) + a_{n-3}) \cdots a_0$$

The properties of polynomials are following

- The polynomial is uniquely identified by its coefficients e.g. $\sum\limits_{k=0}^{n} a_k x^k = \sum\limits_{k=0}^{n} b_k x^k$ iff $a_k = b_k$.
- A polynomial of order $n \geq 1$ has up to n real roots. The only n'th order polynomial with more then n roots is a zero polynomial.
- There is unique n'th order polynomial that intersects with $n + 1$ points.
- It is very easy to compute, integrate or differentiate polynomials.

1.2 Polynomial Interpolation (PI)

Definition 1.3. Let the data known by $y_k = f(x_k)$ be given by a table $(x_k, y_k)_{k=0}^n$. The Polynomial Interpolation for f at points $\{x_k\}_{k=0}^n$ is at least n'th order polynomial $P_n(x)$ that satisfy the **interpolation condition** $P_n(x_k) = y_k$ for all $0 \leq k \leq n$.

There is more then one polynomial that satisfies $P_n(x_k) = y_k$, but there is unique one that has the order $\leq n$.

Theorem 1.4 (Existence & Uniqueness of PI). Let f be a function defined in $[a, b]$, for any set of different points $\{x_k\}_{k=0}^n$ there exists unique polynomial of order $\leq n$.
Proof: Interpolation condition $P_n(x_k) = y_k$ creates linear system of equation for the coefficients a_k, the resulting matrix is Vandermunde, which is non singular as long as $\{x_k\}_{k=0}^n$ are different.

The proof suggests using Vandermonde for interpolation, however Inverting Vandermonde is numerically unstable process and also takes $O(n^3)$ computational steps for $n \times n$ matrix.

1.3 Lagrange Interpolation

Lagrange Polynomial reads for $P_n(x) = \sum\limits_{k=0}^{n} l_{k,n}(x) f(x_k)$ where $l_{k,n}(x) = \frac{\tilde{l}_{k,n}(x)}{\tilde{l}_{k,n}(x_k)} = \delta_{k,i} = \begin{cases} 1 & i = k \\ 0 & else \end{cases}$ and $\tilde{l}_{k,n}(x) = (x - x_0) \cdots (x - x_{k-1})(x - x_{k+1}) \cdots (x - x_n)$.
Note that $\tilde{l}_{k,n}(x_j) = 0$ for $j \neq k$, which gives

$$P_n(x_j) = \sum_{k=0}^{n} l_{k,n}(x_j) f(x_k) = \frac{\tilde{l}_{j,n}(x_j)}{\tilde{l}_{j,n}(x_j)} f(x_j) = f(x_j)$$

1.4 Newton Interpolation

We would like to construct an interpolation polynomial of the third form.

Proposition 1.5. If

$$P_n(x) = a_0 \cdots + a_k(x - x_0) \cdots (x - x_{k-1}) \cdots \\ +a_n(x - x_0) \cdots (x - x_{n-1})$$

is the polynomial interpolation of function $f(x)$ at points $x_0, \cdots, x_n$ then the polynomial

$$P_n(x) = a_0 + + + a_k(x - x_0) \cdots (x - x_{k-1})$$

is the polynomial interpolation of function $f(x)$ at points $x_0, \cdots, x_k$.

Corollary 1.6. If P_k and P_{k-1} are polynomial interpolation of $f(x)$ at points $x_0, ..., x_k$, $x_0, ..., x_{k-1}$ respectively. Then

$$P_k(x) = P_{k-1}(x) + a_k(x - x_0) \cdots (x - x_{k-1}).$$

Furthermore, a_j depends only on $x_0, ..., x_j$, but not on $x_{j+1}, ..., x_k$.

Definition 1.7 (Divided Differences (DD)). Consider set of points $\{x_k\}_{k=0}^n$. For each point x_k define $f[x_k] = f(x_k)$. For any set of points $x_0, ..., x_k$ with $x_k \neq x_0$ the divided differences are defined as

$$f[x_0, ..., x_k] = \frac{f[x_1, ..., x_k] - f[x_0, ..., x_{k-1}]}{x_k - x_0}$$

Proposition 1.8.

$$P_n(x) = f[x_0] + f[x_0, x_1](x - x_0) + + + \\ f[x_0, ..., x_n](x - x_0) \cdots (x - x_{n-1})$$

In other words, $a_0 = f[x_0]$, $a_1 = f[x_0, x_1]$, $a_k = f[x_0, x_1, ..., x_k]$ for $0 \leq k \leq n$.

Proposition 1.9. $f[x_0, ..., x_k]$ is independent of the order of the points.
Proof: Since $f[x_0, ..., x_k]$ is the leading coefficient of the polynomial interpolation. Since the polynomial interpolation is independent of the order of the points and unique, its leading coefficient should also be independent of the order of the points.

The convenient way to construct DD is the triangular table as following (the required coefficient underlined):

$$\begin{array}{lll}
x_0 \ \underline{f[x_0]} & & \\
 & \underline{f[x_0, x_1] = \frac{f[x_1] - f[x_0]}{x_1 - x_0}} & \\
x_1 \ f[x_1] & & \underline{f[x_0, x_1, x_2] = \cdots} \\
 & f[x_1, x_2] = \frac{f[x_2] - f[x_1]}{x_2 - x_1} & \\
x_2 \ f[x_2] & &
\end{array}$$

1.5 Adding a point

Assume we add a new point (x_{n+1}, y_{n+1}) to P_n. The additional work to do in direct (Vandermonde), Lagrange and Newton's methods is:

1. Direct: Recalculate from the start, i.e. $\mathcal{O}(n^3)$.
2. Lagrange: multiply each $l_i(x)$ by $\frac{x - x_{n+1}}{x_i - x_{n+1}}$ and calculate $l_{n+1}(x)$. That's $\mathcal{O}(n^2)$ operations.
3. Newton: Add a line to the triangle - $\mathcal{O}(n)$.

1.6 Interpolation Error

Theorem 1.10. Let $f(x)$ be defined in $[a, b]$ and let $P_n(x)$ be PI of $f(x)$ at points $\{x_k\}_{k=0}^n \subset [a, b]$, then the interpolation error for $x \in [a, b]$ is given by

$$e(\tilde{x}) = f(x) - P_n(x) = f[x_0, \ldots, x_n, \tilde{x}] \prod_{k=0}^{n} (\tilde{x} - x_k)$$

Proof: Since $P_n(x)$ is polynomial interpolation of $f(x)$, we have $e(\tilde{x}_k) = f(x_k) - P_n(x_k) = 0$ for any of $\{x_k\}_{k=0}^n$. Let $\tilde{x} \notin \{x_k\}_{k=0}^n$ be new interpolation point and let construct new interpolation point $P_{n+1}(x) = P_n(x) + f[x_0, \ldots, x_n, \tilde{x}] \prod_{j=0}^{n} (x - x_j)$ therefore $e(\tilde{x}) = f(\tilde{x}) - P_n(\tilde{x}) = P_{n+1}(\tilde{x}) - P_n(\tilde{x}) = f[x_0, \ldots, x_n, \tilde{x}] \prod_{j=0}^{n} (\tilde{x} - x_j)$

Definition 1.11. A polynomial is called a monic polynomial if its leading coefficient (the nonzero coefficient of highest degree) is equal to 1.

Proposition 1.12. $(x - x_0) \cdots (x - x_n)$ is monic.

Corollary 1.13. $P_n(x) = f[x_0, ..., x_n]x^n + Q_{n-1}(x)$.

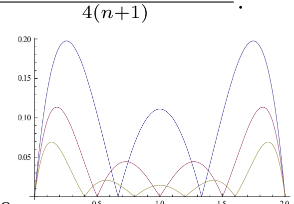
Theorem 1.14. Let $f(x) \in \mathcal{C}^n$ in $[a, b]$ and let $\{x_k\}_{k=0}^n \subset [a, b]$. There is exists $c \in [a, b]$ such that

$$f[x_0, ..., x_n] = \frac{f^{(n)}(c)}{n!}$$

Corollary 1.15. $e(x) = \frac{f^{(n+1)}(c)}{(n+1)!} \prod_{k=0}^{n} (x - x_k)$

One bounds the error by $|e(x)| \leqslant \frac{\max_{x_0 \leqslant x \leqslant x_n} |f^{(n+1)}(x)|}{(n+1)!} |p(x)|$ Note the two factors of the error.

1. $|f^{(n+1)}|$ - uncontrollable and may be affected round off error etc.
2. $p(x) = \prod_{k=0}^{n} |x - x_k| \leqslant |b - a|^{n+1}$. The later approximation is coarse. For the better one we need more information about $p(x)$ or x_k. The common choice is uniform grid , i.e. $x_j = x_0 + jh$, $h = \frac{b-a}{n}$ then $p(x)$ is bounded by $\frac{h^{n+1}}{4} n!$ and $|e(x)| \leqslant \frac{h^{n+1} \max_{x_0 \leqslant c \leqslant x_n} |f^{(n+1)}(c)|}{4(n+1)}$.



The error grows near the ends of the interval with uniform grid. Thus, stay away from the ends, i.e. work inside the interval. See the figure of $\prod_{j=0}^{n} |x - j\frac{2-0}{n}|$ on $[a, b] = [0, 2]$ for $n \in \{3, 4, 5\}$.

Conclusion: It is recommended to avoid using interpolation by high order polynomials.