Lagrange Polynomials, Differentiation and Richardson Extrapolation

---

1. (a) By definition $p_{i,i}(x) = f(x_i)$ is a polynomial of degree zero (a constant) which takes the value $f(x_i)$ when $x = x_i$. So that satisfies the requirements.

   To show the second bit we use induction. It's a bit fiddly. So assume the result is true for j: that is $p_{i,j}(x)$ is a polynomial of degree at most $j - i$ and $p_{i,j}(x_k) = f(x_k)$ for all $0 \le i \le k \le j \le n$. Then

$$p_{i,j+1}(x) = \frac{(x - x_{j+1})p_{i,j}(x) - (x - x_i)p_{i+1,j+1}(x)}{x_i - x_{j+1}}$$

   is obviously a polynomial of degree at most $j + 1 - i$ (one degree higher because of multiplying by x). Also

$$p_{i,j+1}(x_k) = \frac{(x_k - x_{j+1})p_{i,j}(x_k) - (x_k - x_i)p_{i+1,j+1}(x_k)}{x_i - x_{j+1}}.$$

   We can use assumptions $p_{i,j}(x_k) = f(x_k)$ for $0 \le i \le k \le j$ and $p_{i+1,j+1}(x_k) = f(x_k)$ for $i + 1 \le k \le j + 1 \le n$ to see that for $i + 1 \le k \le j$

$$p_{i,j+1}(x_k) = \frac{(x_k - x_{j+1})f(x_k) - (x_k - x_i)f(x_k)}{x_i - x_{j+1}} = f(x_k).$$

   Also, when $0 \le i = k \le j + 1$,

$$p_{i,j+1}(x_i) = \frac{(x_i - x_{j+1})p_{i,j}(x_i) - (x_i - x_i)p_{i+1,j+1}(x_i)}{x_i - x_{j+1}} = f(x_i)$$

   (because the $x_i - x_i = 0$ which kills off the term $p_{i+1,j+1}(x_i)$ whose value cannot be assumed) and when $i \le k = j + 1$ we have

$$p_{i,j+1}(x_{j+1}) = \frac{(x_{j+1} - x_{j+1})p_{i,j}(x_{j+1}) - (x_{j+1} - x_i)p_{i+1,j+1}(x_{j+1})}{x_i - x_{j+1}} = f(x_{j+1})$$

   for a similar reasoning. So the result is also true for $j + 1$ and since it is true when $j = i$, then it is true for all $i \le j \le n$.

   (b) The algorithm goes like this. First step: define $p_{i,i}(x) = f(x_i)$ for all $0 \le i \le n$. Next step: define $p_{i,i+1}(x)$ using $j = i + 1$ in the given formula for $0 \le i \le n - 1$ which only require the values of $p_{i,i}(x)$ which are known. Then we set $j = i + 2$ for $0 \le i \le n - 2$ to define $p_{i,i+2}(x)$ and so on until the last step when we set $j = i + n$ for $i = 0$ only to define $p_{0,n}(x)$ which is a polynomial of degree at most n which equals $f(x_k)$ for $0 \le k \le n$.

   Hence $p_{0,n}(x) = P_n(x)$, since this is the *unique* Lagrange interpolating polynomial.

(c) Let's try and estimate the number of flops (floating point operations) needed for the algorithm above. These are any of the operations $+, -, *, /$ needed to make the computation of $P_n(x)$ for a given value of x. The first step counts nothing. There are n iterations thereafter. The ith step involves definining $n-i$ new functions $p_{i,j}(x)$ which involves some $+, -, *, /$ operations (let's not bother counting them). So the whole process of determining $p_{0,n}(x)$ is roughly $O((n-1)+(n-2)+...+2+1) = O((n-1)(n-2)/2)$ operations.

Now the Lagrange polynomial interpolation formula. This is the sum over $n+1$ values of j of the product over n factors. So this is approximately $O(n(n+1))$ operations.

So both are $O(n^2)$ operations, but the prefactor of 1/2 in the method in (b) means it is roughly double the speed (all other things being equal – one really does need to count all of the operations needed at each step to be precise).

Note: The method in this question is called Neville's algorithm[1].

2. To confirm, substitute in. So $y'(x) = 2ax + b$ and the RHS is

$$\frac{a(x+h)^2 + b(x+h) + c - a(x-h)^2 - b(x-h) - c}{2h} = 2ax + b.$$

And again, $y''(x) = 2a$ and the RHS is

$$\frac{a(x+h)^2 + b(x+h) + c - 2ax^2 - 2bx - 2c + a(x-h)^2 + b(x-h) + c}{2h} = 2a.$$

Why ? Well the central difference scheme for the derivative and the formula for the second derivative both have errors proportional to $y'''(\xi)$ and here $y'''(x) \equiv 0$.

3. (a) Using Taylor's theorem we find

$$f(x_0 + h) = f(x_0) + h\, f'(x_0) + \frac{h^2}{2} f''(\xi)\,,$$

where $\xi \in (x_0, x_0 + h)$. We solve this equation for $f'(x_0)$ and obtain

$$f'(x_0) = \frac{f(x_0 + h) - f(x_0)}{h} - \frac{h}{2} f''(\xi)\,.$$

This shows that the truncation error is $-hf''(\xi)/2$. The approximation is of order h and is a first order scheme. Since $f(x) = \mathrm{e}^x$ the maximum value of $|f''(\xi)|\, \exp^{1+h}$ and so a bound on the error introduced when calculating $f'(1)$ is $(h/2)\exp^{1+h}$.

(b) We apply the approximation to evaluate the derivative of $f(x) = \mathrm{e}^x$ at $x = 1$. The following table shows the results of a numerical evaluation of $f'(1) \approx [f(1+h) - f(1)]/h$ with four decimal places accuracy for various values of h. The best result is obtained for $h = 10^{-2}$.

[1] after E.H. Neville, an English mathematician who convinced the more famous mathematician Ramanujan to come to England from India

h	$f'(1)$	error E
10^{-0}	4.6708	1.9525
10^{-1}	2.8590	0.1407
10^{-2}	2.7300	0.0117
10^{-3}	2.7000	-0.0183
10^{-4}	3.0000	0.2817
10^{-5}	0.0000	-2.7183

(c) We can estimate the optimal value, h_{opt}, of h theoretically. Denote the round-off errors for $f(1+h)$ and $f(1)$ are denoted by $e(1+h)$ and $e(1)$. Then, for four digit precision, machine accuracy is $\epsilon = 10^{-4}$ and the round-off errors are bounded by $\epsilon|f(1+h)| = 10^{-4}|f(1+h)| \lesssim 10^{-4}\mathrm{e}^1$ for h small and $\epsilon|f(1)| = 10^{-4}|f(1)| = 10^{-4}\mathrm{e}^1$

From part (a) the truncation error is bounded by $(h/2)\mathrm{e}^{1+h} \approx (h/2)\mathrm{e}$.

This yields the following bound for the total error E_t

$$|E_t| = \left|\frac{e(1+h)-e(1)}{h} - \frac{h}{2}f''(\xi)\right| \leq \left|\frac{e(1+h)}{h}\right| + \left|\frac{e(1)}{h}\right| + \left|\frac{h}{2}f''(\xi)\right| \leq \frac{2\epsilon\mathrm{e}}{h} + \frac{\mathrm{e}h}{2}.$$

This bound is minimal (setting $d|E_t|/dh = 0$) when $h = h_{opt} = 2\sqrt{\epsilon} \approx 0.02$, in good agreement with the numerically obtained optimal value of $h \approx 10^{-2}$.

4. (a) There is just one backward difference approximation of $f'(1) \approx 0.2$ and a forward difference of $f'(1) \approx 0.8$.

   (b) It's clear the hint is useful as we have evaluations of f at $x_0 = 1$ and $x_0 - h$, $x_0 + 2h$ where $h = 0.1$. We approximate $f'(x)$ by

$$\begin{aligned} \alpha f(x_0-h) + \beta f(x_0) + \gamma f(x_0+2h) &= \alpha(f(x_0) - hf'(x_0) + \tfrac{1}{2}h^2 f''(x_0) - \tfrac{1}{6}h^3 f'''(\xi_1)) \\ &\quad +\beta f(x_0) \\ &\quad +\gamma(f(x_0) + 2hf'(x_0) + \tfrac{1}{2}(2h)^2 f''(x_0) + \tfrac{1}{6}(2h)^3 f'''(\xi_2)) \end{aligned}$$

where $\xi_1 \in (x_0 - h, x_0)$ and $\xi_2 \in (x_0, x_0 + 2h)$. Equating terms proportional to $f(x_0)$ and setting the resulting coefficient to zero gives

$$\alpha + \beta + \gamma = 0.$$

Equating terms proportional to $f'(x_0)$ and setting the resulting coefficient to 1 gives

$$2\gamma - \alpha = 1/h$$

and for $f''(x_0)$ we want the resulting coefficient to be zero so

$$4\gamma + \alpha = 0.$$

We solve these to get $\alpha = -2/3h$, $\gamma = 1/6h$ and $\beta = 1/2h$ and so this tells us that

$$f'(x_0) = \frac{-4f(x_0 - h) + 3f(x_0) + f(x_0 + 2h)}{6h} + E(h)$$

where

$$E(h) = \frac{h^3}{6}(\alpha f'''(\xi_1) - 8\gamma f'''(\xi_2)) = -\frac{h^2}{9}(f'''(\xi_1) + 2f'''(\xi_2))$$

is $O(h^2)$. If we use the numbers supplied we get

$$f'(1) \approx \frac{-3.92 + 3 + 1.16}{0.6} = 0.4.$$

(c) With $f(x) = 1 + 20(x-1)^3$ we have $f'(1) = 0$ and therefore the 1st order backward difference approximation of 0.2 is better than the 2nd order scheme giving 0.4. Why is this ? Well, the truncation error in the backward difference scheme is, according to the notes,

$$\frac{1}{2}hf''(\xi), \qquad \text{for } \xi \in (0.9, 1).$$

Here, $h = 0.1$, $f''(x) = 120(x-1)$ and so $-f''(\xi)$ takes a minimum value of 0 and maximum value of 12. Therefore the truncation error lies somewhere between 0 and $12 \times 0.1 \times \frac{1}{2} = -0.6$.

For the second order scheme, the truncation error is, according to part (b),

$$E(h) = -\frac{1}{9}h^2(f'''(\xi_1) + 2f'''(\xi_2))$$

and $h^2 = 0.01$ but $f'''(x) = 120$. So we have

$$E(h) = -\frac{1}{9} \times 0.01 \times 120 \times 3 = -0.4$$

precisely. The error is large here because $f'''(x)$ is large even though h^2 is small. However, $f''(x)$ may not be as large which is why the first order scheme happens to outperform the second order scheme.

(d) If a quadratic function is fitted to the data then the approximatin in (b) is exact since the error is proportional to third derivatives and these are zero for quadratics. There would still be an error in 1st order schemes since the error is proportional to $f''(x)$ which is not, in general, zero.

5. (a) The forward difference and central difference approximations for the derivative $f'(1.2)$ for $h = 0.1$ and $h = 0.2$ are

$$f'(1.2) \approx \frac{f(1.3) - f(1.2)}{0.1} = \frac{3.535581 - 3.094479}{0.1} = 4.4110200$$

$$f'(1.2) \approx \frac{f(1.4) - f(1.2)}{0.2} = \frac{3.996196 - 3.094479}{0.2} = 4.5085850$$

$$f'(1.2) \approx \frac{f(1.3) - f(1.1)}{0.2} = \frac{3.535581 - 2.677335}{0.2} = 4.2912300 = f'_h$$

$$f'(1.2) \approx \frac{f(1.4) - f(1.0)}{0.4} = \frac{3.996196 - 2.287355}{0.4} = 4.2721025 = f'_{2h}$$

(b) Using Taylor's theorem we find that

$$
\begin{aligned}
&af(x_0) + bf(x_0+h) + cf(x_0+2h) \\
&= af(x_0) + b\left[f(x_0) + hf'(x_0) + \frac{h^2}{2}f''(x_0) + \frac{h^3}{6}f'''(\xi_1)\right] \\
&\quad + c\left[f(x_0) + 2hf'(x_0) + 2h^2 f''(x_0) + \frac{4h^3}{3}f'''(\xi_2)\right] \\
&= (a+b+c)f(x_0) + h(b+2c)f'(x_0) + \frac{h^2}{2}(b+4c)f''(x_0) + \frac{h^3}{6}(bf'''(\xi_1) + 8cf'''(\xi_2)),
\end{aligned}
$$

where $\xi_1 \in (x_0, x_0+h)$ and $\xi_2 \in (x_0, x_0+2h)$. To obtain an approximation for $f'(x_0)$ we require

$$a+b+c=0,\ h(b+2c)=1,\ b+4c=0 \implies c=-1/2h,\ b=2/h,\ a=-3/2h.$$

This results in the following approximation

$$f'(x_0) = \frac{-3f(x_0) + 4f(x_0+h) - f(x_0+2h)}{2h} + O(h^2).$$

A numerical evaluation with $x_0 = 1.2$ and $h = 0.1$ results in

$$f'(1.2) \approx \frac{-3 \times 3.094479 + 4 \times 3.535581 - 3.996196}{0.2} = 4.313455.$$

(c) We know from the lecture that the error of the central difference approximation is of order h^2. If we denote the third approximation in the solutions of part (a) by f'_h we have (with $h = 0.1$)

$$f'(x_0) = f'_h + Ah^2 + O(h^3), \tag{1}$$

where A is some constant. (In fact we could write $O(h^4)$ instead of $O(h^3)$ because we showed in the lecture that the expansion of the error term involves only even powers of h.) If we denote the fourth approximation in the solutions of part (a) by f'_{2h} we have (with $h = 0.1$)

$$f'(x_0) = f'_{2h} + 4Ah^2 + O(h^3). \tag{2}$$

We can get rid of the h^2-term by forming the combination $[4 \times (1) - (2)]/3$

$$f'(x_0) = \frac{4f'_h - f'_{2h}}{3} + O(h^3).$$

A numerical evaluation yields

$$f'(x_0) \approx \frac{4 \times 4.2912300 - 4.2721025}{3} = 4.297606.$$

This is a much better approximation than all previous ones. The actual value is $f'(1.2) = 4.297549$.

6. The magnitude of the total error at $x = 1$ is

$$|E_t| \approx \left| \frac{\epsilon(f(1+h) - 2f(1) + f(1-h))}{h^2} - \frac{1}{12}h^2 f^{(iv)}(\xi) \right|$$

where $\xi \in (1-h, 1+h)$ and ϵ is machine accuracy. Thus

$$|E_t| \lesssim \frac{4\epsilon e^{-1}}{h^2} + \frac{1}{12}h^2|f^{(iv)}(1)|$$

since $f(1 \pm h) \approx f(1) = e^{-1}$. We need $4e^{-x^2}(3 - 12x^2 + 4x^4)$ which gives $|f^{(iv)}(1)| = 20e^{-1}$ so that

$$|E_t| \lesssim \frac{4\epsilon e^{-1}}{h^2} + \frac{5}{3}h^2 e^{-1}$$

Minimising as a function of h gives

$$h_{opt} = (12\epsilon/5)^{1/4} \approx 0.012$$

7. We need to consider the full Taylor series expansion of the approximation to $f''(x_0)$, thus

$$\begin{aligned}\frac{f(x_0+h) - 2f(x_0) + f(x_0-h)}{h^2} &= (1/h^2)[(f + hf' + (h^2/2)f'' + \ldots + (h^k/k!)f^{(k)} + \ldots) \\ &\quad -2f \\ &\quad +(f - hf' + (h^2/2)f'' - \ldots + (-1)^k(h^k/k!)f^{(k)} + \ldots)]\end{aligned}$$

(omitting the argument x_0 from the function f for brevity). And when we tidy everything up we see that the odd powers of h cancel and we get

$$f''(x_0) = \frac{f(x_0+h) - 2f(x_0) + f(x_0-h)}{h^2} - 2\sum_{k=1}^{\infty} \frac{h^{2k}}{(2k+2)!} f^{(2k+2)}(x_0)$$

Thus, we may write $N = f''$, $N_1(h)$ is the approximation to N and the equation above takes the form

$$N = N_1(h) + a_2h^2 + a_4h^2 + \ldots.$$

Let $h \to h/2$ to get a new approximation

$$N = N_1(h/2) + a_2h^2/4 + a_4h^2/16 + \ldots$$

It follows that

$$N = N_2(h) + O(h^4)$$

when

$$N_2(h) = \frac{4N_1(h/2) - N_1(h)}{3}$$

In terms of the original evaluations of f, $N_2(h)$ is

$$\frac{4}{3}\frac{f(x_0+h/2) - 2f(x_0) + f(x_0-h/2)}{(h/2)^2} - \frac{1}{3}\frac{f(x_0+h) - 2f(x_0) + f(x_0-h)}{h^2}$$

which simplifies to the expression in the question and is evidently $O(h^4)$ accurate.

8. We apply Taylor's theorem to the following linear combination

$$
\begin{aligned}
&\alpha f(x_0) + \beta f'(x_0 + h) + \gamma f(x_0 - \lambda h) \\
&= \alpha f(x_0) + \beta \left[f'(x_0) + h f''(x_0) + \frac{h^2}{2} f'''(x_0) + \frac{h^3}{6} f^{(iv)}(\xi_1) \right] \\
&\quad + \gamma \left[f(x_0) - \lambda h f'(x_0) + \frac{\lambda^2 h^2}{2} f''(x_0) - \frac{\lambda^3 h^3}{6} f'''(x_0) + \frac{\lambda^4 h^4}{24} f^{(iv)}(\xi_2) \right] \\
&= (\alpha + \gamma) f(x_0) + (\beta - \lambda h \gamma) f'(x_0) + \frac{h}{2}(2\beta + \lambda^2 h \gamma) f''(x_0) + \frac{h^2}{6}(3\beta - \lambda^3 h \gamma) f'''(x_0) \\
&\quad + \frac{h^3}{24}(4\beta f^{(iv)}(\xi_1) + \lambda^4 h \gamma f^{(iv)}(\xi_2)\,,
\end{aligned}
$$

where ξ_1 is between x_0 and $x_0 + h$, and ξ_2 is between x_0 and $x_0 - \lambda h$. To obtain an approximation for $f''(x_0)$ we require

$$
\alpha + \gamma = 0, \qquad \beta - \lambda h \gamma = 0, \qquad \frac{h}{2}(2\beta + \lambda^2 h \gamma) = 1.
$$

This leads to the following values

$$
\gamma = \frac{2}{\lambda h^2 (2 + \lambda)}, \qquad \beta = \frac{2}{h(2 + \lambda)}, \qquad \alpha = \frac{-2}{\lambda h^2 (2 + \lambda)}.
$$

The resulting approximation is

$$
f''(x_0) = \frac{-2f(x_0) + 2\lambda h f'(x_0 + h) + 2f(x_0 - \lambda h)}{\lambda h^2 (2 + \lambda)} + \frac{(3 - \lambda^2)h}{3(2 + \lambda)} f'''(x_0) + O(h^2).
$$

We see that the error is in general $O(h)$. However, if $\lambda^2 = 3$ then the term proportional to h vanishes and we obtain an error of order $O(h^2)$.