Contents

1 Polynomial Interpolation

Let $x_0, x_1, \cdots, x_n$ and $y_0, y_1, \cdots, y_n$ $(n \in \mathbb{N})$ be given real or complex numbers where the x_i's, called *nodes* or abscissae, are *distinct* (no two the same). A function g *interpolates* the data if it satisfies

$$g(x_i) = y_i \qquad i = 0 \to n.$$

Lagrange Interpolation

THEOREM. 1.1 *(Existence and Uniqueness) Given $x_0, \cdots, x_n$ and $y_0, \cdots, y_n$ real or complex numbers, where the x_i's are distinct. Then there exists a unique $p_n \in \mathcal{P}_n :=$* {polynomials of degree less than or equal to n} *such that*

$$p_n(x_i) = y_i \qquad i = 0 \to n.$$

Uniqueness is important as there are many ways of writing the interpolating polynomial, e.g. Lagrange, Newton's divided difference and backward/forward difference formulae, but all give the same polynomial.

Truncation Error of Lagrange Interpolation

Let $p_n \in \mathcal{P}_n$ interpolate the function f at $n+1$ distinct nodes x_i, i.e.

$$p_n(x) = \sum_{i=0}^{n} l_i(x) f(x_i) \quad \text{where} \quad l_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^{n} \frac{x - x_j}{x_i - x_j}$$

and let the *truncation error* be defined to be

$$E_n(x) = f(x) - p_n(x).$$

Notice we can rewrite

$$l_i(x) = \frac{w_{n+1}(x)}{(x - x_i) w'_{n+1}(x_i)} \quad \text{where} \quad w_{n+1}(x) := \prod_{j=0}^{n} (x - x_j).$$

EXAMPLE. The *Chebyshev polynomials*

$$T_n(x) := \cos n\theta \qquad \text{where } \theta = \cos^{-1} x \tag{1.1}$$

satisfies $T_0(x) = 1$, $T_1(x) = x$ and the following three term recurrence relation

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \quad n = 1, 2, \cdots$$

from which it follows that $T_n(x) = 2^{n-1}x^n + \cdots$ for $n = 1, 2, \cdots$.

THEOREM. 1.2 *Let $\{x_0,\cdots,x_n\}\subset[a,b]$ be distinct nodes, $f\in C^{n+1}[a,b]$ and $p_n\in\mathcal{P}_n$ interpolate f at the nodes. Then for all $x\in[a,b]$*

$$E_n(x):=f(x)-p_n(x)=\frac{w_{n+1}(x)}{(n+1)!}f^{(n+1)}(\xi)\qquad\text{where } \xi=\xi(x)\in(a,b). \tag{1.2}$$

Hermite Interpolation

At distinct nodes $\{x_i\}_{i=0}^n$ fit a polynomial with heights $\{y_i\}_{i=0}^n$ and slopes $\{y_i'\}_{i=0}^n$.

THEOREM. 1.3 *The polynomial* $p_{2n+1}(x)=\sum_{i=0}^n\left[h_i(x)y_i+\overline{h}_i(x)y_i'\right]\in\mathcal{P}_{2n+1}$ *where*

$$\overline{h}_i(x):=(x-x_i)l_i^2(x),\quad h_i(x):=[1-2(x-x_i)l_i'(x_i)]\,l_i^2(x).$$

has the property $p_{2n+1}(x_i)=y_i$ *and* $p_{2n+1}'(x_i)=y_i'$ *for* $i=0\to n$, *since*

$$h_i(x_j)=\delta_{ij},\ h_i'(x_j)=0,\ \overline{h}_i(x_j)=0 \text{ and } \overline{h}_i'(x_j)=\delta_{ij}.$$

We call this *osculatory interpolation.* It is often referred to as *Hermite interpolation* although we reserve this term for the more general case:

THEOREM. 1.4 *Let $\{x_i\}_{i=0}^n$ be distinct real (or complex) numbers and $f^{(j)}(x_i)$ $(0\leqslant j\leqslant k_i)$ $(0\leqslant i\leqslant n)$ be given. Then there is a unique polynomial $p_N\in\mathcal{P}_N$ where $N=\sum_{i=0}^n(k_i+1)-1$ such that $p_N^{(j)}(x_i)=f^{(j)}(x_i)$ $(0\leqslant j\leqslant k_i)$ $(0\leqslant i\leqslant n)$.*

THEOREM. 1.5 *(Truncation Error) Let $f\in C^{2n+2}[a,b]$ and p_{2n+1} be the interpolating osculatory polynomial at the distinct nodes x_i $(i=0\to n)$. Then for all $x\in[a,b]$*

$$E(x):=f(x)-p_{2n+1}(x)=\frac{[w_{n+1}(x)]^2}{(2n+2)!}f^{(2n+2)}(\xi)\qquad\xi\in(a,b) \tag{1.3}$$

In summary, polynomials are inappropriate for general approximation of functions because a sequence of interpolating polynomials will *not* always converge as n increases and even if it does it may have to be of high degree to be accurate.

THEOREM. 1.6 *Weierstrass approximation. Given $f\in C[a,b]$ and $\varepsilon>0$ there exists $n=n(\varepsilon)\in\mathbb{N}$ and $p_n\in\mathcal{P}_n$ such that*

$$|f(x)-p_n(x)|<\varepsilon\quad\forall\ x\in[a,b].$$

PROOF. It is sufficient to prove the Weierstrass theorem when $[a,b]=[0,1]$.

For any $n\in\mathbb{N}$ and $k=0\to n$ define

$$\beta_{n,k}(x):=\binom{n}{k}x^k(1-x)^{n-k}\in\mathcal{P}_n\quad\text{and}\quad B_n(f;x):=\sum_{k=0}^n\beta_{n,k}(x)f(\tfrac{k}{n})\in\mathcal{P}_n$$

where $x\in[0,1]$; we call $B_n(f;x)$ the *Bernstein polynomial.* $B_n(f;0)=f(0)$ and $B_n(f;1)=f(1)$ but in general $B_n(f;\frac{k}{n})\neq f(\frac{k}{n})$. From the following identities

$$(x+y)^n=\sum_{k=0}^n\binom{n}{k}x^ky^{n-k},\quad n(x+y)^{n-1}=\sum_{k=0}^n\binom{n}{k}kx^{k-1}y^{n-k},\quad n(n-1)(x+y)^{n-2}=\sum_{k=0}^n\binom{n}{k}k(k-1)x^{k-2}y^{n-k}$$

by setting $y=1-x$ we deduce the first three identities and the fourth follows from the definition of $\beta_{n,k}$.

$$\text{a) } 1=\sum_{k=0}^n\beta_{n,k}(x),\quad\text{b) } nx=\sum_{k=0}^nk\beta_{n,k}(x),\quad\text{c) } n(n-1)x^2=\sum_{k=0}^nk(k-1)\beta_{n,k}(x),\quad\text{d) } \beta_{n,k}(x)\geqslant0. \tag{1.4}$$

We now show that $|f(x)-B_n(f;x)|<\varepsilon$. Given $\varepsilon>0$ there exists $\delta=\delta(\varepsilon)$ such that $|f(x)-f(\frac{k}{n})|<\frac{\varepsilon}{2}$ for all $x\in[0,1]$ where $|x-\frac{k}{n}|<\delta$. Noting (1.4a)

$$\begin{aligned}f(x)-B_n(f;x)&=\sum_{k=0}^n\beta_{n,k}(x)(f(x)-f(\tfrac{k}{n}))\\&=\sum_{S_1(x)}\beta_{n,k}(x)(f(x)-f(\tfrac{k}{n}))+\sum_{S_2(x)}\beta_{n,k}(x)(f(x)-f(\tfrac{k}{n}))\end{aligned} \tag{1.5}$$

where $S_1(x):=\{k\in\{0,1,\ldots,n\}:|x-\frac{k}{n}|<\delta\}$ and $S_2(x):=\{k\in\{0,1,\ldots,n\}:|x-\frac{k}{n}|\geqslant\delta\}$. Noting (1.4a,d)

$$|\sum_{S_1(x)}\beta_{n,k}(x)(f(x)-f(\tfrac{k}{n}))|\leqslant\sum_{S_1(x)}\beta_{n,k}(x)|f(x)-f(\tfrac{k}{n})|<\tfrac{\varepsilon}{2}. \tag{1.6}$$

Since f is continuous $|f(x)|\leqslant M$ for all $x\in[0,1]$ it follows on noting (1.4a-d) and $x\in[0,1]$ that

$$\begin{aligned}|\sum_{S_2(x)}\beta_{n,k}(x)(f(x)-f(\tfrac{k}{n}))|&\leqslant2M\sum_{S_2(x)}\beta_{n,k}(x)\leqslant2M\sum_{S_2(x)}\tfrac{(x-\frac{k}{n})^2}{\delta^2}\beta_{n,k}(x)\\&\leqslant\tfrac{2M}{n^2\delta^2}\sum_{k=0}^n(nx-k)^2\beta_{n,k}(x)=\tfrac{2M}{n^2\delta^2}\underbrace{(n^2x^2-2nx(nx)+n(n-1)x^2+nx)}_{=nx(1-x)}\leqslant\tfrac{M}{2n\delta^2}\leqslant\tfrac{\varepsilon}{2}\end{aligned} \tag{1.7}$$

where we choose $n\geqslant M/(\delta^2\varepsilon)$. Hence the results follows by combining (1.5–1.7).

2 Piecewise polynomial approximation

In this chapter we consider piecewise (broken) polynomial approximations, it turns out that they are much more satisfactory.

Piecewise linear interpolation

Let $\pi : a = x_0 < x_1 < \cdots < x_n = b$ be a *partition* of the interval $[a,b] \subset \mathbb{R}$ upon which we wish to approximate a continuous function f. We will call $\{x_i\}_{i=0}^n$ "knots". The continuous, piecewise linear, function, s, interpolating f is defined by

$$s(x) = \frac{x_{j+1}-x}{x_{j+1}-x_j} f(x_j) + \frac{x-x_j}{x_{j+1}-x_j} f(x_{j+1}) \quad x \in [x_i, x_{i+1}] \ (i = 0 \to n-1)$$

We can rewrite s as

$$s(x) = \sum_{j=0}^n \lambda_i(x) f(x_i)$$

where

$$\lambda_0(x) = \begin{cases} \frac{x_1-x}{x_1-x_0} & x_0 \leqslant x \leqslant x_1 \\ 0 & x_1 \leqslant x \leqslant x_n \end{cases}, \quad \lambda_i(x) = \begin{cases} \frac{x-x_{i-1}}{x_i-x_{i-1}} & x_{i-1} \leqslant x \leqslant x_i \\ \frac{x_{i+1}-x}{x_{i+1}-x_i} & x_i \leqslant x \leqslant x_{i+1} \\ 0 & \text{otherwise} \end{cases}, \quad \lambda_n(x) = \begin{cases} 0 & x_0 \leqslant x \leqslant x_{n-1} \\ \frac{x-x_{n-1}}{x_n-x_{n-1}} & x_{n-1} \leqslant x \leqslant x_n. \end{cases}$$

Notice that if $f \in C^2[a,b]$, then from linear interpolation for $x \in [x_i, x_{i+1}]$,

$$|f(x) - s(x)| \leqslant \frac{(x_{i+1}-x_i)^2}{8} \|f''\|_\infty,$$

where $\|f''\|_\infty := \max_{x\in[a,b]} |f''(x)|$, hence for all $x \in [a,b]$

$$|f(x) - s(x)| \leqslant \frac{h^2}{8} \|f''\|_\infty \qquad \text{where } h = \max_i x_{i+1} - x_i.$$

Least-squares approximation by piecewise linear functions.

Let $L_1 f$ be the least-squares approximation to f by a piecewise linear polynomial. That is $L_1 f(x) = \sum_{i=0}^n \alpha_i \lambda_i(x)$, where the coefficients α_i are chosen so as to minimize

$$E = \int_a^b [f(x) - L_1 f(x)]^2 \, \mathrm{d}x.$$

The coefficients α_i which minimize E satisfy $\partial E / \partial \alpha_i = 0$, i.e.

$$\sum_{j=0}^n \alpha_j \int_a^b \lambda_i(x)\lambda_j(x) \mathrm{d}x = \int_a^b \lambda_i(x) f(x) \mathrm{d}x. \qquad (2.1)$$

As $\int_a^b \lambda_i(x)\lambda_j(x)\mathrm{d}x = 0$ when $|i-j| > 1$, we can easily calculate (2.1). Let $h_{i-1} := x_i - x_{i-1}$, then for $i = 1 \to n-1$ we have

$$\int_a^b \lambda_i^2(x)\mathrm{d}x = \frac{h_{i-1}+h_i}{3} \text{ and } \int_a^b \lambda_i(x)\lambda_{i-1}(x)\mathrm{d}x = \frac{h_{i-1}}{6}.$$

So that multiplying (2.1) by $6/(h_{i-1}+h_i)$ yields

$$\frac{h_{i-1}}{h_{i-1}+h_i}\alpha_{i-1} + 2\alpha_i + \frac{h_i}{h_{i-1}+h_i}\alpha_{i+1} = 3\beta_i \qquad (i = 1 \to n-1)$$

$$\text{where} \quad \beta_i := \frac{2}{h_{i-1}+h_i} \int_{x_{i-1}}^{x_{i+1}} \lambda_i(x) f(x) \mathrm{d}x.$$

Likewise, performing a similar calculation when $i = 0$ and n, it results from equation (2.1)

$$2\alpha_0 + \alpha_1 = 3\beta_0 := \frac{6}{h_0}\int_{x_0}^{x_1} \lambda_0(x) f(x)\mathrm{d}x, \qquad \alpha_{n-1} + 2\alpha_n = 3\beta_n := \frac{6}{h_{n-1}} \int_{x_{n-1}}^{x_n} \lambda_n(x) f(x) \mathrm{d}x.$$

Notice that $\|\beta_i\|_\infty \leqslant \|f\|_\infty$.

THEOREM. 2.1 *(Diagonally dominant matrices) Let A be a square matrix with elements a_{ij} $(i,j = 1 \to n)$. If A is strictly diagonally dominant, i.e. $\sum_{\substack{j=1 \\ j\neq i}}^n |a_{ij}| < |a_{ii}|$ then A is non-singular.*

Hence, we see that $L_1 f$ exists and is unique.

A bound on the error.

Let $s(x)$ be the piecewise linear interpolant of f on π

$$\|f - L_1 f\|_\infty = \|f - s + s - L_1 f\|_\infty \leqslant \|f - s\|_\infty + \|L_1 s - L_1 f\|_\infty = \|f - s\|_\infty + \|L_1(s - f)\|_\infty.$$

Can we bound $\|L_1 g\|_\infty$ in terms of $\|g\|_\infty$? Notice that $\|L_1 g\|_\infty = \max_{i=0\to n} |\alpha_i| = |\alpha_j|$ and assuming $1 \leqslant j \leqslant n-1$, from the j'th equation

$$2\|L_1 g\|_\infty = |2\alpha_j| = \left| 3\beta_j - \frac{\alpha_{j-1}h_{j-1} + \alpha_{j+1}h_j}{h_{j-1}+h_j} \right| \leqslant 3|\beta_j| + |\alpha_j| \leqslant 3\|g\|_\infty + \|L_1 g\|_\infty$$

(If $j = 0$ or n the same inequality holds trivially). Therefore assuming that $f \in C^2[a,b]$

$$\|f - L_1 f\|_\infty \leqslant 4\|f - s\|_\infty \leqslant \frac{h^2}{2}\|f''\|_\infty.$$

Piecewise cubic interpolation

If f is smoother, we may be able to get a better approximation more efficiently by looking at piecewise polynomials of higher degree.

Method 1 On $[x_i, x_{i+1}]$ $(i = 1 \to n-2)$ use the polynomial which interpolates f at $x_{i-1}, x_i, x_{i+1}, x_{i+2}$. On $[x_0, x_1]$ we use the the first four nodes and on $[x_{n-1}, x_n]$ we use the last four nodes. If s is the resulting approximation and $f \in C^4[a,b]$ then for all $x \in [a,b]$

$$|f(x) - s(x)| \leqslant \frac{h^4}{16}\|f^{(4)}\|_\infty \qquad \text{where} \quad h = \max x_{i+1} - x_i$$

which follows from the Lagrange interpolation truncation error theorem.

Method 2 On $[x_i, x_{i+1}]$ construct s so that

$$s(x_i) = f(x_i),\ s'(x_i) = s_i' \qquad (i = 0 \to n)$$

and s is continuously differentiable (s_i' is unknown). A calculation gives

$$s(x) = f(x_i) + s_i'(x - x_i) + c_i^2(x - x_i)^2 + c_i^3(x - x_i)^3 \quad x \in [x_i, x_{i+1}]$$

where

$$c_i^2 = \frac{3(f(x_{i+1}) - f(x_i))}{h_i^2} - \frac{2s_i' + s_{i+1}'}{h_i}, \quad c_i^3 = \frac{-2(f(x_{i+1}) - f(x_i))}{h_i^3} + \frac{s_{i+1}' + s_i'}{h_i^2}$$

If $s'(x_i) = f'(x_i)$ $(i = 0 \to n)$ this is *piecewise cubic Osculatory interpolation* on $[x_i, x_{i+1}]$ and if $f \in C^4[a, b]$ then the following error equation holds

$$|f(x) - s(x)| \leqslant \left|\frac{(x - x_i)^2(x - x_{i+1})^2}{4!} f^{(4)}(\xi)\right| \leqslant \left(\frac{h^2}{4}\right)^2 \frac{\|f^{(4)}\|_\infty}{4!} = \frac{h^4}{384}\|f^{(4)}\|_\infty.$$

where $h = \max x_{i+1} - x_i$.

Each of these approximations is "local" (*only* what happens a short distance from the interval is important) making the solution easy to generate. A non-local $C^2[a, b]$ piecewise cubic approximation is:

Cubic Splines

Let $\pi : a = x_0 < x_1 < \cdots < x_n = b$ be a given partition of the real interval $[a, b]$. A cubic spline with knots $x_0, x_1, \cdots, x_n$ is a function s which has the following properties:

1. $s(x)$ is a cubic polynomial in each interval (x_i, x_{i+1}) $(i = 0 \to n - 1)$.
2. $s \in C^2[a, b]$.

We now construct an interpolating cubic spline, s, explicitly. It is convenient to introduce $M_i := s''(x_i)$ and $f_i := f(x_i)$ $(i = 0 \to n)$. As s is piecewise cubic, in the interval (x_i, x_{i+1}), $s''(x)$ is linear, and for $i = 0 \to n - 1$

$$s''(x) = \frac{x - x_i}{h_i} M_{i+1} + \frac{x_{i+1} - x}{h_i} M_i,$$

where $h_i = x_{i+1} - x_i$. Integrating twice and applying the continuity conditions $s(x_i) = f_i$ and $s(x_{i+1}) = f_{i+1}$, we find that that for $x \in (x_i, x_{i+1})$ $(i = 0 \to n - 1)$,

$$s(x) = \frac{(x - x_i)^3}{6h_i} M_{i+1} + \frac{(x_{i+1} - x)^3}{6h_i} M_i + \left(f_i - \frac{h_i^2}{6} M_i\right) \frac{x_{i+1} - x}{h_i} + \left(f_{i+1} - \frac{h_i^2}{6} M_{i+1}\right) \frac{x - x_i}{h_i}.$$

For s' to be continuous on $[a, b]$, s' has to be continuous at the internal knots, i.e.

$$\lim_{x \to x_i^+} s'(x) = \lim_{x \to x_i^-} s'(x) \qquad (i = 1 \to n - 1).$$

$$\lim_{x \to x_i^+} s'(x) = -\tfrac{h_i}{3} M_i - \tfrac{h_i}{6} M_{i+1} + \tfrac{f_{i+1} - f_i}{h_i} \tag{2.2}$$

$$\lim_{x \to x_{i+1}^-} s'(x) = \tfrac{h_i}{3} M_{i+1} + \tfrac{h_i}{6} M_i + \tfrac{f_{i+1} - f_i}{h_i} \tag{2.3}$$

Hence, matching equations (2.2) and (2.3) for $i = 1 \to n - 1$ and multiplying by $\frac{6}{h_i + h_{i-1}}$ yields

$$\mu_i M_{i-1} + 2M_i + \lambda_i M_{i+1} = d_i \qquad i = 1 \to n - 1.$$

where

$$\mu_i = h_{i-1}/(h_i + h_{i-1}),\ \lambda_i = 1 - \mu_i = h_i/(h_i + h_{i-1}) \text{ and } d_i = \frac{6}{h_i + h_{i-1}}\left[\frac{f_{i+1} - f_i}{h_i} - \frac{f_i - f_{i-1}}{h_{i-1}}\right].$$

Notice that with **uniform knot spacing** $x_i = a + ih$ where $h = (b - a)/n$, we have

$$M_{i-1} + 4M_i + M_{i+1} = \frac{6}{h^2}\Delta^2 f_{i-1} \qquad (\text{where } \Delta f_i = f_{i+1} - f_i).$$

We need to assign two extra conditions. Which we will analyse.

End conditions

1. *Natural cubic spline*. Let $s''(a) = 0 = s''(b)$, i.e. $M_0 = 0 = M_n$, the system we must solve is

$$A\boldsymbol{m} = \begin{pmatrix} 2 & \lambda_1 & 0 & \cdots & 0 \\ \mu_2 & 2 & \lambda_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \mu_{n-2} & 2 & \lambda_{n-2} \\ 0 & \cdots & 0 & \mu_{n-1} & 2 \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \\ \vdots \\ M_{n-2} \\ M_{n-1} \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_{n-2} \\ d_{n-1} \end{pmatrix} = \boldsymbol{d}$$

where $|\lambda_1|, |\mu_{n-1}| < 1$ and $|\mu_i| + |\lambda_i| = 1 < 2$ $(i = 2 \to n - 2)$. That is the matrix is strictly diagonally dominant, and hence non-singular. So for given values $\{f_i\}$ there exists a unique natural cubic spline interpolating the data at the knots.

2. *Complete cubic spline* (clamped or displine)

If we know $f'(x_0)$ and $f'(x_n)$ we could take

$$\lim_{x \to x_0^+} s'(x) = f'(x_0) \quad \text{and} \quad \lim_{x \to x_n^-} s'(x) = f'(x_n).$$

From (2.2) with $i = 0$ and (2.3) with $i = n - 1$

$$f'(x_0) = \frac{-h_0}{3} M_0 - \frac{h_0}{6} M_1 + \frac{f_1 - f_0}{h_0} \iff 2M_0 + M_1 = \frac{6}{h_0}\left[\frac{f_1 - f_0}{h_0} - f'(x_0)\right]$$

$$f'(x_n) = \frac{h_{n-1}}{3} M_n + \frac{h_{n-1}}{6} M_{n-1} + \frac{f_n - f_{n-1}}{h_{n-1}} \iff 2M_n + M_{n-1} = \frac{6}{h_{n-1}}\left[f'(x_n) - \frac{f_n - f_{n-1}}{h_{n-1}}\right].$$

From strict diagonal dominance there is a unique solution in this case too.

Alternatively, we could specify $s'(x_0)$ and $s'(x_n)$ to be arbitrary numbers.

3. Generalize 1. and 2. to

$$2M_0 + \lambda_0 M_1 = d_0 \quad \text{and} \quad \mu_n M_{n-1} + 2M_n = d_n \quad \text{where } d_0, d_n \text{ are arbitrary.}$$

   If $|\lambda_0|, |\mu_n| < 2$ then again from strict diagonal dominance of the matrix there exists a unique solution.

4. *Not-a-knot*

   We could specify that s''' is continuous at x_1 and x_{n-1} so

$$\frac{M_1 - M_0}{h_0} = \lim_{x\to x_1^-} s'''(x) = \lim_{x\to x_1^+} s'''(x) = \frac{M_2 - M_1}{h_1} \quad \text{and} \quad \frac{M_n - M_{n-1}}{h_{n-1}} = \frac{M_{n-1} - M_{n-2}}{h_{n-2}}$$

   or

$$-\mu_1 M_0 + M_1 - \lambda_1 M_2 = 0 \quad \text{and} \quad -\mu_{n-1} M_{n-2} + M_{n-1} - \lambda_{n-1} M_n = 0.$$

   Using the existing equations for M_0, M_1 and M_2 to eliminate M_0 and similarly eliminate M_n to obtain a system which is strictly diagonally dominant.

5. *Parabolic runout* We set the end conditions to be $M_0 = M_1$ and $M_{n-1} = M_n$.

Of the end-conditions mentioned above, only the complete cubic spline guarantees that the piecewise cubic spline of itself is itself.

Spline functions of degree k

A *spline function of degree* k on the partition $\pi : a = x_0 < x_1 < \cdots < x_n = b$ is a function s_k which has the following properties

(i) $s_k(x)$ is a polynomial of degree k on each interval $[x_i, x_{i+1}]$ $(i = 0 \to n-1)$.

(ii) $s_k(x) \in C^{k-1}[a,b]$.

Any spline function of degree k, with knots $x_0, \cdots, x_n$ may be expressed as

$$s_k(x) = \sum_{i=0}^{k} c_i x^i + \frac{1}{k!}\sum_{j=1}^{n-1} g_j (x - x_j)_+^k \qquad a\leqslant x\leqslant b$$

where the *truncated power functions* are

$$(x - x_j)_+^k = \begin{cases} (x-x_j)^k & \text{if } x > x_j \\ 0 & \text{if } x\leqslant x_j \end{cases} \qquad k\geqslant 0,$$

g_j is the jump discontinuity in the k'th derivative at x_j, i.e. $g_j = \lim_{x\to x_j^+} s_k^{(k)}(x) - \lim_{x\to x_j^-} s_k^{(k)}(x)$ and $\sum_{i=0}^k c_i x^i$ is the spline defined on $[x_0, x_1]$. Notice the above function is $k-1$ times continuously differentiable and the number of unknowns is $k+1+n-1$.

This formulation may be unsatisfactory since there might be a huge loss of accuracy through cancellation.

B-splines

A spline of degree 0 with support $(x_i, x_{i+1}]$ is

$$B_i^0(x) = \begin{cases} 1, & \text{if } x_i < x\leqslant x_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

define

$$B_i^1(x) := \frac{x - x_i}{x_{i+1} - x_i} B_i^0(x) + \frac{x_{i+2} - x}{x_{i+2} - x_{i+1}} B_{i+1}^0(x) \Longrightarrow B_i^1(x) = \begin{cases} \frac{x - x_i}{x_{i+1} - x_i} & \text{if } x_i < x\leqslant x_{i+1} \\ \frac{x_{i+2} - x}{x_{i+2} - x_{i+1}} & \text{if } x_{i+1} < x\leqslant x_{i+2} \\ 0 & \text{otherwise.} \end{cases}$$

In fact $B_i^1(x) \equiv \lambda_{i+1}(x)$, in old hat function notation, is a spline function of degree 1 with support $[x_i, x_{i+2}]$. Now defining the *B-spline of degree* k, B_i^k, recursively for each i and $k = 1, \cdots$

$$B_i^k(x) := \frac{x - x_i}{x_{i+k} - x_i} B_i^{k-1}(x) + \frac{x_{i+k+1} - x}{x_{i+k+1} - x_{i+1}} B_{i+1}^{k-1}(x)$$

it is clearly a piecewise positive, polynomial of degree k with support $[x_i, x_{i+k+1}]$.

THEOREM. 2.2 *$B_i^k(x)$ is a spline function of degree k with support $[x_i, x_{i+k+1}]$.*

PROOF. Embed $x_0, \cdots, x_n$ in infinite knot sequence

$$\cdots < x_{-2} < x_{-1} < x_0 < \cdots < x_n < x_{n+1} < \cdots.$$

For equally spaced knots the recurrence relation becomes

$$B_i^k(x) := \frac{x - x_i}{kh} B_i^{k-1}(x) + \frac{x_{i+k+1} - x}{kh} B_{i+1}^{k-1}(x) = \frac{1}{kh}\left[(x - x_i) B_i^{k-1}(x) + (x_{i+k+1} - x) B_{i+1}^{k-1}(x)\right]$$

Next we show that for $k > 1$ and for all x

$$\frac{\mathrm{d}}{\mathrm{d}x} B_i^k(x) = \frac{1}{h}\left[B_i^{k-1}(x) - B_{i+1}^{k-1}(x)\right]$$

For $k = 1$ the above equation is true except at the knots. Assume the statement to be true for all i. Differentiating the B-splines recurrence relation, substituting in the hypothesis and rearranging

$$\begin{aligned}
\tfrac{\mathrm{d}}{\mathrm{d}x} B_i^k(x) &= \tfrac{1}{kh}\left[B_i^{k-1}(x) - B_{i+1}^{k-1}(x) + (x - x_i)\tfrac{\mathrm{d}}{\mathrm{d}x} B_i^{k-1}(x) + (x_{i+k+1} - x)\tfrac{\mathrm{d}}{\mathrm{d}x} B_{i+1}^{k-1}(x)\right] \\
&= \tfrac{1}{kh}\left[B_i^{k-1}(x) - B_{i+1}^{k-1}(x) + \tfrac{x - x_i}{h}\left[B_i^{k-2}(x) - B_{i+1}^{k-2}(x)\right] + \tfrac{x_{i+k+1} - x}{h}\left[B_{i+1}^{k-2}(x) - B_{i+2}^{k-2}(x)\right]\right] \\
&= \tfrac{1}{kh}\left[B_i^{k-1}(x) - B_{i+1}^{k-1}(x) + (k-1) B_i^{k-1}(x) + B_{i+1}^{k-2}(x) - (k-1) B_{i+1}^{k-1}(x) - B_{i+1}^{k-2}(x)\right] \\
&= \tfrac{1}{h}\left[B_i^{k-1}(x) - B_{i+1}^{k-1}(x)\right]
\end{aligned}$$

This formula holds everywhere except possibly at the knots. However noting that B_i^1 is continuous for all i it follows by induction that $B_i^k \in C^{k-1}(-\infty, \infty)$ i.e. B_i^k is a spline function of degree k.

One can use the B-splines as a basis for the spline function space. The advantages of expressing the spline in terms of B-splines are: only a few multiplications are required to calculate the value of the splines; one does not suffer from severe cancellation; the splines have small support; calculating the B-spline from the recurrence relation is a computationally stable process.

For example instance let $s(x) = \sum_{i=-3}^{n-1} \alpha_i B_i^3(x)$ be the natural interpolating cubic splines. Noting that

$$B_i^3(x) = \frac{1}{6h^3}\begin{cases} (x-x_i)^3 & x_i \leqslant x \leqslant x_{i+1} \\ h^3 + 3h^2(x-x_{i+1}) + 3h(x-x_{i+1})^2 - 3(x-x_{i+1})^3 & x_{i+1} \leqslant x \leqslant x_{i+2} \\ h^3 + 3h^2(x_{i+3}-x) + 3h(x_{i+3}-x)^2 - 3(x_{i+3}-x)^3 & x_{i+2} \leqslant x \leqslant x_{i+3} \\ (x_{i+4}-x)^3 & x_{i+3} \leqslant x \leqslant x_{i+4} \\ 0 & \text{otherwise} \end{cases}$$

and

$$(B_i^3(x))'' = \frac{1}{6h^3}\begin{cases} 6(x-x_i) & x_i \leqslant x \leqslant x_{i+1} \\ 6h - 18(x-x_{i+1}) & x_{i+1} \leqslant x \leqslant x_{i+2} \\ 6h - 18(x_{i+3}-x) & x_{i+2} \leqslant x \leqslant x_{i+3} \\ 6(x_{i+4}-x) & x_{i+3} \leqslant x \leqslant x_{i+4} \\ 0 & \text{otherwise} \end{cases}$$

noting that

x	x_i	x_{i+1}	x_{i+2}	x_{i+3}	x_{i+4}
$6B_i^3(x)$	0	1	4	1	0
$h^2(B_i^3(x))''$	0	1	-2	1	0

we find that interpolating and natural boundary conitions correspond to:

$$f(x_k) = \sum_{i=-3}^{n-1} \alpha_i B_i^3(x_k) = \frac{1}{6}(\alpha_{k-3} + 4\alpha_{k-2} + \alpha_{k-1}) \quad (k = 0 \to n)$$

$$0 = h^2 \sum_{i=-3}^{n-1} \alpha_i (B_i^3(x_0))'' = \alpha_{-3} - 2\alpha_{-2} + \alpha_{-1} \quad \text{and} \quad \alpha_{n-3} - 2\alpha_{n-2} + \alpha_{n-1} = 0$$

thus eliminating α_{-3} and α_{n-1} we get the following strictly diagonally dominant system of linear equations to solve

$$\begin{pmatrix} 6 & 0 & 0 & \cdots & 0 \\ 1 & 4 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & 4 & 1 \\ 0 & \cdots & 0 & 0 & 6 \end{pmatrix} \begin{pmatrix} \alpha_{-2} \\ \alpha_{-1} \\ \alpha_0 \\ \vdots \\ \alpha_{n-2} \end{pmatrix} = 6 \begin{pmatrix} f(x_0) \\ f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{pmatrix}$$

That is the cubic B-splines form a basis for the natural cubic splines.

It is possible to prove a similar result to the Weierstrass theorem, see Powell's book on Approximation theory and methods.

THEOREM. 2.3 *(Holladay's Theorem.) Let $\pi : a = x_0 < x_1 < \cdots < x_n = b$ be a partition and $y \in C^2[a,b]$ be arbitrary such that y interpolates a function f at the knots. The natural cubic spline, s, with those knots uniquely minimizes*

$$\int_a^b y''(x)^2 \, dx.$$

In other words, the natural cubic spline gives the smallest value for the integral over the admissible class of functions.

PROOF. Let $y \in C^2[a,b]$ and $y(x_j) = f(x_j)$ $(j = 0 \to n)$. Let s be the interpolating natural cubic spline with knots $x_0, \cdots, x_n$, i.e. $s(x_j) = f(x_j)$ $(j = 0 \to n)$ and $s''(a) = s''(b) = 0$. From the identity $\alpha^2 - \beta^2 = (\alpha-\beta)^2 + 2\beta(\alpha-\beta)$, we get

$$\int_a^b y''(x)^2 dx - \int_a^b s''(x)^2 dx = \int_a^b (y''(x) - s''(x))^2 \, dx + 2\int_a^b s''(x)\,(y''(x) - s''(x))\, dx.$$

The last integral is zero; this is called *the first integral relation.* Now it is clear that

$$\int_a^b y''(x)^2 dx \geqslant \int_a^b s''(x)^2 dx$$

with equality if and only if

$$\int_a^b (y''(x) - s''(x))^2 \, dx = 0.$$

That is $y''(x) \equiv s''(x)$ hence $y(x) = s(x) + cx + d$ and applying $s(a) = y(a)$ and $s(b) = y(b)$ we find that $\int_a^b y''(x)^2 dx$ is minimized if and only if $y(x) = s(x)$. □

COROLLARY 2.4 *Let $y \in C^2[a,b]$ be arbitrary, such that y interpolates the function f at the knots x_j $(j = 0 \to n)$ and satisfies the additional conditions $y'(a) = f'(a)$ and $y'(b) = f'(b)$. The complete cubic spline, s, with those knots uniquely minimizes $\int_a^b y''(x)^2 dx$.*

THEOREM. 2.5 *(Error) Let $f \in C^2[a,b]$, then the second derivative s'' of the complete cubic spline S is the least-squares approximation to f'' by piecewise linear functions on the partition π. Furthermore, if $f \in C^4[a,b]$, then for all $x \in [a,b]$,*

$$\|f^{(i)}(x) - s^{(i)}(x)\|_\infty \leqslant C_i h^{4-i} \|f^{(4)}\|_\infty, \ (i = 0, 1, 2)$$

where $h = \max x_{j+1} - x_j$ and $f^{(i)} = d^{(i)} f / dx^{(i)}$.

3 Minimax and near-minimax approximation

Introduction

Any "measure of distance" of an approximating function to a given function will suffice to give us an error. The natural setting for *approximation theory* is a metric space, however normed linear spaces have nicer properties. A normed linear space *is* a metric space, where $d(x,y) = \|x-y\|$.

Let X be a vector space over a field $\mathbb{F}$ and $x \in X$. The *norm* of x is a non-negative number, $\|x\|$, with the properties

1. $\|x\| \geqslant 0 \ \forall\, x \in X$ and $\|x\| = 0$ iff $x = 0$.
2. $\|cx\| = |c|\|x\| \ \forall\, x \in X$ and $\forall\, c \in \mathbb{F}$.
3. $\|x+y\| \leqslant \|x\| + \|y\| \ \forall\, x, y \in X$.

Let $C[a,b]$, as usual be all those continuous, real-valued functions on the interval $[a,b]$. For $1 \leqslant p < \infty$ and $f \in C[a,b] = X$ we define the L^p-norm to have the value

$$\|f\|_p = \left[\int_a^b |f(x)|^p \mathrm{d}x\right]^{1/p}.$$

The L^∞-norm (also called uniform norm) has the value

$$\|f\|_\infty = \max_{a \leqslant x \leqslant b} |f(x)|.$$

Only $(X, \|.\|_\infty)$ gives us a complete metric space.

For discrete problems, when $X = \mathbb{R}^m$, we can define analogous norms, i.e. for $\boldsymbol{g} \in \mathbb{R}^m$

$$\|\boldsymbol{g}\|_\infty = \max_{1 \leqslant i \leqslant m} |g_i| \qquad \text{and} \qquad \|\boldsymbol{g}\|_p = \left[\sum_{i=1}^m |g_i|^p\right]^{1/p} \quad 1 \leqslant p < \infty.$$

Best Approximation

Given $\{\phi_i\}_{i=0}^n$, an appropriate basis, and the function $f(x)$ on the interval $[a,b]$ we consider the approximations of the form

$$F_n(x) = \sum_{i=0}^n a_i \phi_i(x).$$

A *best approximation* is one which minimizes the distance. In general, different norms yield *different* best approximations.

A *minimax approximation* minimizes

$$\|f - F_n\|_\infty := \max_{x \in [a,b]} |f(x) - F_n(x)|.$$

A *least-squares approximation* minimizes

$$\|f - F_n\|_2 = \left[\int_a^b |f(x) - F_n(x)|^2 \mathrm{d}x\right]^{1/2}.$$

One can also consider approximations using discrete norms.

EXAMPLE. On $[0,1]$ find some best constant approximations to e^x.

Given f and $\|\cdot\|$ we can ask:

Is there a best approximation? $\exists$	How can it be found? Algorithm.
Is it unique? !	What is the error like? Accuracy
Can it be describe? Characterization	How does the accuracy change with n? Convergence.

THEOREM. 3.1 *Let $E_n(f) := \min_{a_i} \|f(x) - \sum_{i=0}^n a_i \phi(x)\|$ for the given function and chosen norm. If F_n and G_n are two best approximations, then for $\alpha \in (0,1)$, $\alpha F_n + (1-\alpha) G_n$ is also a best approximation. Furthermore $E_n(f)$ does not increase with n.*

THEOREM. 3.2 *(De la Vallée Poussins) Given $f \in C[a,b]$ and $n \geqslant 0$. Suppose $q_n(x) \in \mathcal{P}_n$ such that*

$$f(x_j) - q_n(x_j) = (-1)^j e_j \qquad (j = 0 \to n+1)$$

where $e_j \neq 0$ and all of the same sign, $a \leqslant x_0 < \cdots < x_{n+1} \leqslant b$. Then

$$\min_{j=0 \to n+1} |e_j| \leqslant \|f - p_n^*\|_\infty \leqslant \|f - q_n\|_\infty$$

where p_n^* is the the *minimax polynomial of degree n for f on $[a,b]$.*

THEOREM. 3.3 *(Equioscillation) For a given function $f \in C[a,b]$, a polynomial p_n is a minimax polynomial of degree n if and only if there is a sequence of $n+2$ points $\{x_j\}$, with $a \leqslant x_0 < x_1 < \cdots < x_{n+1} \leqslant b$, at which $E_n(x) = f(x) - p_n(x)$ takes alternately the values $\pm E_n$ where $E_n := \|f - p_n\|_\infty$.*

PROOF. (a) **sufficiency** This follows de la Vallée Poussins theorem.

(b) **necessity**. Suppose that p_n is a minimax approximation of degree n for f on the interval $I = [a,b]$. Let $E_n(x) := f(x) - p_n(x)$. If $E_n(x) \equiv 0$ on I there is nothing to prove. Otherwise, suppose that $E_n := \max_{x \in I} |E_n(x)|$ is attained at $\eta_0, \cdots, \eta_r$ $(r \geqslant 1)$, where the η_j are well ordered, i.e. $a \leqslant \eta_0 < \eta_1 < \cdots < \eta_r \leqslant b$. Then $E_n(\eta_j) = \pm E_n$. Consider the sign changes of $E_n(\eta_j)$ as j increases from 0 to r. If the k'th change occurs between η_i and η_{i+1}, let $\alpha_k = \frac{1}{2}(\eta_i + \eta_{i+1})$. Assume there are s such that points α_k $(k = 1 \to s)$. There exists a unique polynomial q_s of degree s such that $q_s(\alpha_k) = 0$ $(k = 1 \to s)$ and $q_s(\eta_0) = E_n(\eta_0)$. Since $q_s(\eta_j)$ has the same sign as $E_n(\eta_j)$ for $j = 0 \to r$, it follows that for some β sufficiently small, for all $x \in [a,b]$, $|E_n(x) - \beta q_s(x)|$ is smaller than E_n. This means that $\max_{x \in I} \left| f(x) - [p_n(x) + \beta q_s(x)] \right| \leqslant E_n$, i.e. $p_n + \beta q_s$ has a smaller maximum value error than p_n. But p_n is a minimax polynomial of degree n, so $s \geqslant n+1$ (otherwise we contradict the minimax property). From the set $\{\eta_0, \eta_1, \cdots, \eta_r\}$ choose a subset $\{x_0, x_1, \cdots, x_s\}$ with $a \leqslant x_0 < x_1 < \cdots < x_s \leqslant b$, such that $E_n(x_{j+1}) = -E_n(x_j)$ $(j = 0 \to s-1)$. Since $s \geqslant n+1$ there are at least $n+2$ points $x_0, x_1, \cdots, x_{n+1}$. $\square$

EXAMPLES. The minimax polynomial of degree n for x^{n+1} on $[-1,1]$ is p_n^* where $p_n^*(x) = x^{n+1} - \frac{1}{2^n}T_{n+1}(x)$ which follows from the equioscillation property of $T_{n+1}(x)$.

See a previous example for the constant minimax polynomial of e^x on [0,1].

COROLLARY 3.4 *Let p_n^* be a minimax polynomial of degree n on $[a,b]$ for a given function $f \in C^{n+1}[a,b]$. Then there are $n+1$ distinct points $t_0, t_1, \cdots, t_n$ in $[a,b]$ such that*

$$f(x) - p_n^*(x) = \frac{(x-t_0)(x-t_1)\cdots(x-t_n)}{(n+1)!} f^{(n+1)}(\zeta) \quad \text{for some } \zeta \in (a,b).$$

THEOREM. 3.5 *(Uniqueness) Let $f \in C[a,b]$ be given, then there exists at most one minimax polynomial*

THEOREM. 3.6 *(Convergence) Let p_n^* be the minimax polynomial approximation to $f \in C[a,b]$. Then $E_n := \|f - p_n^*\|_\infty \to 0$ as $n \to \infty$.*

Computing the minimax approximation

We call a set of $n+2$ points such that $a \leqslant \xi_0 < \xi_1 < \cdots < \xi_{n+1} \leqslant b$ a *reference*.

THEOREM. 3.7 *(Discrete De la Vallée Poussins) Given a reference $\{\xi_i\}_{i=0}^{n+1} \subset \{x_i\}_{i=0}^m$ $(m > n \geqslant 0)$ and $\{f(x_i)\}_{i=0}^m$, suppose $q_n(x) \in \mathcal{P}_n$ such that*

$$f(\xi_i) - q_n(\xi_i) = (-1)^i e_i \qquad (i = 0 \to n+1)$$

where $e_i \neq 0$ are all of the same sign. Then

$$\min_{i=0\to n+1} |e_i| \leqslant \|f - p_n^*\|_\infty = \max_{i=0\to m} |f(x_i) - p_n^*(x_i)| \leqslant \|f - q_n\|_\infty$$

where p_n^* is the the *minimax polynomial of degree n for f on the discrete point set $\{x_i\}_{i=0}^m$.*

THEOREM. 3.8 *p_n^* is the polynomial of degree n which minimizes the expression*

$$\max_{i=0\to m} |f(x_i) - p_n(x_i)|,$$

over the polynomials of degree n, if and only if there is a reference $\{\xi_i\}_{i=0}^{n+1}$ such that

$$f(\xi_{i+1}) - p_n^*(\xi_{i+1}) = -(f(\xi_i) - p_n^*(\xi_i)) \qquad (i = 0, 1 \to n).$$

PROOF. This is the same as the equioscillation theorem except we replace $[a,b]$ by the x_i' and $\|f - p\|_\infty$ by $\max_{i=0\to m} |f(x_i) - p_n(x_i)|$. □

In principle, calculating the minimax polynomial on a discrete set is easy! Let $p_n^*(x) = \sum_{j=0}^n a_j x^j$, then we have to find $h, \boldsymbol{a}$ such that

$$f(\xi_i) - \sum_{j=0}^n a_j \xi_i^j = (-1)^i h \qquad i = 0 \to n+1;$$

$n+2$ equations for the $n+2$ unknowns.

EXAMPLE. Find the minimax straight line for e^x on the reference $\{0, 0.5, 1\}$, i.e. find $a_0, a_1 \in \mathbb{R}$ such that

$$\max_{x=0,0.5,1} |e^x - (a_0 + a_1 x)|$$

is minimized.

How do we go about calculating the discrete minimax polynomial of degree n where we have $m+1$ points to choose from where $m+1 \geqslant n+2$?

The Discrete Exchange Algorithm

1. The values of a function f are known on a set of point $Z := \{x_i : i = 0 \to m\}$ such that $a \leqslant x_0 < x_1 < \cdots < x_m \leqslant b$. We want to find the polynomial $p_n(x)$ of degree $n < m$ which minimizes $\max_{i=0\to m} |f(x_i) - p_n(x_i)|$

   step 1. Choose a reference $\{\xi_i : i = 0, 1 \to n+1\} \subset Z$.

   step 2. Solve the equation
   $$f(\xi_i) - \sum_{j=0}^n a_j \xi_i^j = (-1)^i h, \qquad i = 0, 1 \to n+1$$
   to find $a_0, a_1, \cdots, a_n$ and h. This gives the polynomial
   $$p_n^{(1)}(x) = \sum_{j=0}^n a_j x^j.$$

   step 3. Calculate $|f(x_i) - p_n^{(1)}(x_i)|$ for the points $x_i \in Z$ which are not in the reference. If none of these exceed $|h|$ then `STOP`.

   step 4. Choose a new reference set $\{\xi_i^+\}$ such that $|f(\xi_i^+) - p_n^{(1)}(\xi_i^+)| \geqslant |h|$ with strict inequality for at least one value of i, and with $\text{sign}\left[f(\xi_{i+1}^+) - p_n^{(1)}(\xi_{i+1}^+)\right] = -\text{sign}\left[f(\xi_i^+) - p_n^{(1)}(\xi_i^+)\right]$

   step 5. Return to step 2.

   There are $\binom{m+1}{n+2}$ different choices for the reference, and for each reference a unique polynomial exists which minimizes $\max |f(\xi_k) - p_n(\xi_k)|$. From the discrete version of the de la Vallée Poussin theorem, the algorithm must strictly increase h at each step, so we must stop in a finite number of steps.

   EXAMPLE. What is the quadratic discrete minimax polynomial of e^x on the set $\{0, 0.2, 0.4, 0.6, 0.8, 1\}$?

2. CONTINUOUS VERSION. To find the minimax polynomial $p_n^*(x)$ of degree n for a function $f \in C[a,b]$.

steps 1. & 2.	As for the discrete case.
step 3.	Find the extrema of $f(x)-p_n^{(1)}(x)$ on $[a,b]$. By the de la Vallée Poussin's theorem $\lvert h\rvert \leqslant \|f-p_n^*\|_\infty \leqslant \|f-p_n^{(1)}\|_\infty$ where p_n^* is the minimax polynomial. If $\|f-p_n^{(1)}\|_\infty - \lvert h\rvert$ is sufficiently small STOP.
steps 4. & 5.	This is the same as the discrete case except the points added to the reference set are extrema of $f(x)-p_n^{(1)}(x)$.

EXAMPLE. Previous example revisited.

One way to choose your initial reference is via the *method of forced oscillation* that is take your reference $\xi_k = \cos\frac{(n+1-k)\pi}{n+1}$. For the intervals $[a,b]$ we take the shifted nodes $\xi_k = \frac{1}{2}(a+b)+\frac{1}{2}(b-a)\cos\frac{(n+1-k)\pi}{n+1}$.

EXAMPLE. A cubic approximation for $\sin\left(\frac{\pi x}{2}\right)$ on $[0,1]$.

The first reference consists of the extrema of the Chebyshev polynomial $T_4(x)$ mapped onto $[0,1]$, i.e. the points $\xi_k = \frac{1}{2}[1+\cos\frac{(4-k)\pi}{4}]$. ($\xi_0=0, \xi_1=\frac{1}{2}(1-\frac{1}{\sqrt{2}}), \xi_2=\frac{1}{2}, \xi_3=\frac{1}{2}(1+\frac{1}{\sqrt{2}}), \xi_4=1$). We now seek $p_3^{(1)}(x)=\sum_{j=0}^{3} a_j x^j$ and h so that $f(\xi_k)-\sum_{j=0}^{3} a_j\xi_k^j = (-1)^k h$ $(k=0\to 4)$ with $f(x)=\sin(\frac{\pi x}{2})$

$$\begin{array}{llrrrrcl}
\xi_0: & & -a_0 & & & & = & h\\
\xi_1: & \sin(\frac{\pi\xi_1}{2}) & -a_0 & -a_1\xi_1 & -a_2\xi_1^2 & -a_3\xi_1^3 & = & -h\\
\xi_2: & \sin(\frac{\pi}{4}) & -a_0 & -\frac{1}{2}a_1 & -\frac{1}{4}a_2 & -\frac{1}{8}a_3 & = & h\\
\xi_3: & \sin(\frac{\pi\xi_3}{2}) & -a_0 & -a_1\xi_3 & -a_2\xi_3^2 & -a_3\xi_3^3 & = & -h\\
\xi_4: & \sin(\frac{\pi}{2}) & -a_0 & -a_1 & -a_2 & -a_3 & = & h
\end{array}$$

Eliminate h by adding or subtracting the first equation appropriately

$$\begin{array}{rcrcrcrcl}
2a_0 & + & 0{\cdot}14645a_1 & + & 0{\cdot}02145a_2 & + & 0{\cdot}00314a_3 & = & 0{\cdot}22801\\
 & & 0{\cdot}5a_1 & + & 0{\cdot}25a_2 & + & 0{\cdot}125a_3 & = & 0{\cdot}125\\
2a_0 & + & 0{\cdot}85355a_1 & + & 0{\cdot}72855a_2 & + & 0{\cdot}62186a_3 & = & 0{\cdot}97366\\
 & & a_1 & + & a_2 & + & a_3 & = & 1
\end{array}$$

This gives $p_3^{(1)}(x) = -0{\cdot}0014+1{\cdot}6103x-0{\cdot}1741x^2-0{\cdot}4362x^3$ and $h=-a_0=1{\cdot}4\times10^{-3}$. Let $E(x)=\sin\frac{\pi x}{2}-p_3^{(1)}(x)$. Then $E'(x)=\frac{\pi}{2}\cos\frac{\pi x}{2}-\frac{dp_3^{(1)}}{dx}$ which has zeros (found by for example Newton-Raphson) at 0·1559, 0·522 and 0·8596, the turning points of $E(x)$. We find that $E(0{\cdot}1559)=1{\cdot}314\times10^{-3}$, $E(0{\cdot}522)=1{\cdot}422\times10^{-3}$ and $E(0{\cdot}8596)=-1{\cdot}329\times10^{-3}$. If $p_3^*(x)$ is the minimax polynomial of degree 3, then by de la Vallée Poussin's theorem,

$$1\cdot4\times10^{-3}\leqslant\|f-p_3^*\|_\infty\leqslant 1\cdot422\times10^{-3}=\|f-p_3^{(1)}\|_\infty,$$

so to 2 significant figures $p_3^{(1)}$ and p_3^* have the same maximum error.

Chebyshev Economization of power series

Chebyshev economization is simple. You are given an interval $[a,b]$ and a function $f(x)$. Compute the Maclaurin polynomial of degree n, p_n, and bound the remainder term. Now compute $q_{j-1}(x)=q_j(x)-\alpha_j T_j(x)$ where $q_n(x)=p_n(x)$, T_j is the appropriate Chebyshev polynomial for the interval $[a,b]$ and α_j is chosen so that $q_{j-1}\in\mathcal{P}_{j-1}$.

EXAMPLE. Find some near minimax approximations to e^x on $[-1,1]$.

To apply the minimax idea on an interval $[a,b]$ the appropriate Chebyshev polynomial of degree n is $T_n(\frac{2x-(a+b)}{b-a})$.

EXAMPLE. Find some near minimax approximations e^x but on $[0,1]$.

EXAMPLE. Using Chebyshev economisation find a degree 5 near minimax approximation of $\sin(\pi x/2)$ $[-1,1]$.

Conclusions

1. Minimax polynomial approximation is relevant in achieving the maximum accuracy over all polynomial approximations of a given degree or smallest degree for a given accuracy
2. The minimax puts a bound on the achievable accuracy
3. Near minimax approximations (forced oscillation or Chebyshev economization) are easier to compute and frequently adequate.

4 Padé approximations

We now consider rational approximations. Let f have a convergent Maclaurin series expansion, i.e.

$$f(x)=c_0+c_1x+c_2x^2+\cdots.$$

The $[m/k]$ *Padé approximant* is a rational function

$$R_{mk}(x)=\frac{p_m(x)}{q_k(x)}=\frac{a_0+a_1x+\cdots+a_mx^m}{b_0+b_1x+\cdots+b_kx^k},$$

such that $f(x)-R_{mk}(x)=\mathcal{O}(x^r)$ (r as large as possible). $R_{mk}(x)$ has $m+k+1$ independent variables (we can multiply the numerator and denominator by an arbitrary constant), we hope to make $r=m+k+1$. We assume that $b_0\neq0$ (otherwise R_{mk} is not analytic at 0!) and set $b_0=1$. As f is analytic,

$$f(x)-R_{mk}(x)=\frac{\left(\sum_{i=0}^{\infty}c_ix^i\right)(1+b_1x+\cdots+b_kx^k)-(a_0+a_1x+\cdots+a_mx^m)}{(1+b_1x+\cdots+b_kx^k)}$$

Now choose a_j $(j = 0 \to m)$ and b_j $(j = 1 \to k)$ so that

$$f(x) - R_{mk}(x) = \frac{d_r x^r + \mathcal{O}(x^{r+1})}{(1 + b_1 x + \cdots + b_k x^k)}. \text{ Note that } \frac{d_r x^r}{(1 + b_1 x + \cdots + b_k x^k)}$$

is an estimate of the truncation error.

EXAMPLE. $f(x) = \sqrt{1+2x} = 1 + x - \frac{1}{2}x^2 + \frac{1}{2}x^3 - \frac{5}{8}x^4 + \cdots \; |x| < \frac{1}{2}$. $R_{00}(x) = 1$, $R_{10}(x) = 1 + x$, $R_{20}(x) = 1 + x - \frac{1}{2}x^2$ $(q_0(x) \equiv 1)$ these are just the Taylor polynomials. Let $R_{21}(x) = (a_0 + a_1 x + a_2 x^2)/(1 + b_1 x)$

$$\begin{aligned} q_1(x)(f(x) - R_{21}(x)) &= (1 + x - \frac{1}{2}x^2 + \frac{1}{2}x^3 - \frac{5}{8}x^4 + \cdots)(1 + b_1 x) \\ &\quad -(a_0 + a_1 x + a_2 x^2) = d_4 x^4 + \mathcal{O}(x^5). \end{aligned}$$

$$\left.\begin{array}{lrcl} x^0: & 1 - a_0 = 0 & \Longleftrightarrow & a_0 = 1 \\ x^1: & 1 + b_1 - a_1 = 0 & \Longleftrightarrow & a_1 = 1 + b_1 \\ x^2: & -\frac{1}{2} + b_1 - a_2 = 0 & \Longleftrightarrow & a_2 = b_1 - \frac{1}{2} \\ x^3: & \frac{1}{2} - \frac{1}{2}b_1 = 0 & \Longleftrightarrow & b_1 = 1 \end{array}\right\} \Longrightarrow a_1 = 2,\; a_2 = \frac{1}{2},$$

i.e. $R_{21}(x) = \left(1 + 2x + \frac{1}{2}x^2\right)/\left(1 + x\right)$. This has a singularity at $x = -1$ whereas f has an essential singularity at $x = -1/2$. $d_4 = -5/8 + b_1/2 = -1/8$, so an estimate of the truncation error is $-x^4/(8 + 8x)$. Now to 3 d.p. $f(0.5) = 1.414$, $R_{21}(0.5) = 1.417$ and $p_3(0.5) = 1.438$ (the Taylor cubic). The estimate of the error for $R_{21}(0.5)$ is 5.2×10^{-3}, whereas the actual error is 3×10^{-3}. Let us briefly analyse and compare the number of operations required using alternative formulations:

$$R_{21}(x) = \frac{1 + 2x + 0.5x^2}{1 + x} = \frac{1 + (2 + 0.5x)x}{1 + x}$$

which requires 2 multiplications, 3 additions and 1 division. If we use continued fractions

$$R_{21}(x) = \frac{1 + 2x + 0.5x^2}{1 + x} = 0.5\frac{(x+1)^2 + 2(x+1) - 1}{1 + x} = 0.5\left[(x + 3) - \frac{1}{x + 1}\right]$$

which requires 1 multiplication, 1 division and 3 additions (a subtraction counting as an addition). Note that when using a finite arithmetic, it may be advantageous to use as few arithmetic operations as possible. □

REMARK. We have just considered a "Taylor like" expansion of Padé approximations, one could equally consider Minimax & Interpolating Padé approximations (see Powell – Approximation theory & methods).

Index