Last updated: 21 July 2015
Disclaimer: Links are followed at own risk when viewing this document electronically.

Chapter 5

APPROXIMATION METHODS

5.1 Introduction

In this chapter we investigate two related problems: the approximation of given functions by other, simpler functions, and the fitting of known functions to given data. In the cases studied here, we approximate a continuous function by means of a polynomial.

5.2 Polynomial interpolation

Consider a function $y(x)$ given in the form of the coordinates (x_i, y_i), $i = 0, 1, \ldots, n$. We may approximate the function with a *polynomial* $p_n(x)$ of degree n that is exactly equal to $y(x)$ at the $n + 1$ given points

$$y_i = p_n(x_i) = a_0 + a_1 x_i + a_2 x_i^2 + \cdots + a_n x_i^n. \tag{5.1}$$

We note that there are $n+1$ coefficients a_i as well as $n+1$ points (x_i, y_i). The polynomial in (5.1) is known as an *interpolating polynomial.*

There are $n + 1$ unknowns a_i that must be determined from the $n + 1$ linear equations

$$\begin{aligned} a_0 + a_1 x_0 + \cdots + a_n x_0^n &= p(x_0) = y_0 \\ a_0 + a_1 x_1 + \cdots + a_n x_1^n &= p(x_1) = y_1 \\ &\vdots \\ a_0 + a_1 x_n + \cdots + a_n x_n^n &= p(x_n) = y_n \end{aligned}$$

which can be written compactly in matrix form as

$$\begin{pmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_0 & \cdots & x_0^n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}$$

If $\det \left(x_j^i \right) \neq 0$, these equations have a unique solution. In practice it is difficult and/or inefficient to solve these equations directly, and so we investigate easier and/or faster methods. Firstly, though, we state a few points of general importance.

We may rightly ask whether this interpolating polynomial is unique. Assume that it is not, and thus there are two interpolating polynomials

$$\begin{aligned} p_n(x) &= a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \\ q_n(x) &= b_0 + b_1 x + b_2 x^2 + \cdots + b_n x^n \end{aligned}$$

Then

$$\begin{aligned} Q(x) &= p_n(x) - q_n(x) \\ &= (a_0 - b_0) + (a_1 - b_1)x + (a_2 - b_2)x^2 + \cdots + (a_n - b_n)x^n \end{aligned}$$

is a polynomial of degree n. But $Q(x) = 0$ at the $n+1$ points $\{x_0, x_1, x_2, \ldots, x_n\}$. However, since $Q(x)$ is of degree n, it may only have n roots. This contradiction is resolved only if $Q(x) = 0$, which implies $p_n(x) = q_n(x)$. We conclude that the interpolating polynomial is unique.

Error analysis. We obtain an estimate of the error made when using the polynomial $p_n(x)$ instead of the function $y(x)$. We know that $y(x) = p_n(x)$ at $x_i, i = 0, 1, 2, \ldots, n$. Now consider

$$F(x) = y(x) - p_n(x) - C\prod_{i=0}^{n}(x - x_i). \tag{5.2}$$

Clearly, $F(x) = 0$ at each of the nodes $\{x_0, x_1, \ldots, x_n\}$. Consider any point $x_{n+1} \notin \{x_0, x_1, \ldots, x_n\}$. We choose C such that $F(x_{n+1}) = 0$, which gives

$$C = \frac{y(x_{n+1}) - p_n(x_{n+1})}{\prod_{i=0}^{n}(x_{n+1} - x_i)}. \tag{5.3}$$

Assume $y(x)$, and hence $F(x)$, is continuous. From the constructions (5.2) and (5.3) it follows that $F(x) = 0$ at the $n{+}2$ nodes $\{x_0, x_1, \ldots, x_n, x_{n+1}\}$. Hence, by the Generalized Rolle's Theorem, there exists a point ξ such that

$$F^{(n+1)}(\xi) = 0.$$

If we differentiate (5.2) we obtain

$$\begin{aligned} 0 &= F^{(n+1)}(\xi) \\ &= y^{(n+1)}(\xi) - 0 - C(n+1)! \end{aligned}$$

and so

$$C = \frac{y^{(n+1)}(\xi)}{(n+1)!},$$

which yields, using (5.3),

$$y(x_{n+1}) - p_n(x_{n+1}) = \frac{y^{(n+1)}(\xi)}{(n+1)!}\prod_{i=0}^{n}(x_{n+1} - x_i).$$

But x_{n+1} was chosen arbitrarily, so that this error expression

$$y(x) - p_n(x) = \frac{y^{(n+1)}(\xi(x))}{(n+1)!}\prod_{i=0}^{n}(x - x_i). \tag{5.4}$$

is true for all x. It is easily verified that (5.4) is also true at $\{x_0, x_1, \ldots, x_n\}$.

5.3 Lagrange's method

We now describe a method developed by Lagrange that is a shortcut for determining the coefficients a_i in (5.1).

For each of the points (x_i, y_i), $i = 0, 1, \ldots, n$, we construct an nth degree polynomial that is equal to zero at each of the other points $x_0, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$

$$L_i(x) = A(x - x_0)(x - x_1)\cdots(x - x_{i-1})(x - x_{i+1})\cdots(x - x_n). \tag{5.5}$$

Furthermore, we demand that this polynomial has the value 1 when $x = x_i$

$$1 = A(x_i - x_0)(x_i - x_1)\cdots(x_i - x_i - 1)(x_i - x_{i+1})\cdots(x_i - x_n)$$

from which we obtain

$$A = \frac{1}{(x_i - x_0)(x_i - x_1)\cdots(x_i - x_{i-1})(x_i - x_{i+1})\cdots(x_i - x_n)}$$

so that

$$L_i(x) = \frac{(x - x_0)(x - x_1)\cdots(x - x_{i-1})(x - x_{i+1})\cdots(x - x_n)}{(x_i - x_0)(x_i - x_1)\cdots(x_i - x_i - 1)(x_i - x_{i+1})\cdots(x_i - x_n)}.$$

The nth degree polynomial through the points (x_i, y_i) is then given by

$$p_n(x) = \sum_{i=0}^{n} y_i L_i(x). \tag{5.6}$$

where the $L_i(x)$ are known as the *Lagrange coefficient polynomials*. We may test if this is the required polynomial by substituting $x = x_k$ into equation (5.6)

$$\begin{aligned} p_n(x_k) &= \sum_{i=0}^{n} y_i L_i(x_k) \\ &= y_0 L_0(x_k) + y_1 L_1(x_k) + \cdots + y_k L_k(x_k) + \cdots + y_n L_n(x_k) \\ &= 0 + 0 + \cdots + y_k(1) + \cdots + 0 \\ &= y_k. \end{aligned}$$

Example 5.1. Approximate the sine function with a polynomial that is equal to $\sin x$ at $x = 0, \frac{\pi}{4}, \frac{\pi}{2}$. Determine the true error at $x = \frac{\pi}{6}$ and compare it to an easily calculable upper bound of the error.

The coordinates of the three points that the interpolating polynomial must pass through are shown the following table:

i	x_i	$y_i = \sin x_i$
0	0	0
1	$\frac{\pi}{4}$	$\frac{1}{\sqrt{2}}$
2	$\frac{\pi}{2}$	1

We first determine the polynomials defined in (5.5)

$$\begin{aligned} L_0(x) &= \frac{\left(x - \frac{\pi}{4}\right)\left(x - \frac{\pi}{2}\right)}{\left(0 - \frac{\pi}{4}\right)\left(0 - \frac{\pi}{2}\right)} = \frac{8}{\pi^2}\left(x^2 - \frac{3\pi}{4}x + \frac{\pi^2}{8}\right) \\ L_1(x) &= \frac{(x - 0)\left(x - \frac{\pi}{2}\right)}{\left(\frac{\pi}{4} - 0\right)\left(\frac{\pi}{4} - \frac{\pi}{2}\right)} = -\frac{16}{\pi^2}\left(x^2 - \frac{\pi}{2}x\right) \\ L_2(x) &= \frac{(x - 0)\left(x - \frac{\pi}{4}\right)}{\left(\frac{\pi}{2} - 0\right)\left(\frac{\pi}{2} - \frac{\pi}{4}\right)} = \frac{8}{\pi^2}\left(x^2 - \frac{\pi}{4}x\right) \end{aligned}$$

The interpolating polynomial is given by

$$\begin{aligned} p_2(x) &= y_0 L_0(x) + y_1 L_1(x) + y_2 L_2(x) \\ &= 0 + \frac{1}{\sqrt{2}} L_1 + (1) L_2 \\ &= -\frac{8\left(\sqrt{2} - 1\right)}{\pi^2} x^2 + \frac{2\left(2\sqrt{2} - 1\right)}{\pi} x \end{aligned}$$

We determine the approximation error at $x = \frac{\pi}{6}$

$$\left|\sin\left(\frac{\pi}{6}\right) - p_2\left(\frac{\pi}{6}\right)\right| = |0.5 - 0.517| = 0.017$$

It is interesting to compare this error with the upper limit (5.4)

$$\begin{aligned}\left|\sin\left(\frac{\pi}{6}\right)-p_2\left(\frac{\pi}{6}\right)\right| &\leq \frac{\max_{[0,\pi/2]}|\cos x|}{3!}\left|\prod_{i=0}^{2}\left(\frac{\pi}{2}\right)\right| \\ &= \frac{1}{6}\left|\left(\frac{\pi}{6}-0\right)\left(\frac{\pi}{6}-\frac{\pi}{4}\right)\left(\frac{\pi}{6}-\frac{\pi}{2}\right)\right| \\ &= 0.024\end{aligned}$$

We see that using (5.4) does indeed give, in this case, an upper bound for the magnitude of the approximation error.

5.4 Least-squares curve fitting

A function $y = y(x)$ may be approximated by an nth degree polynomial that passes through each data point, as in §5.2 and §5.3 . However, if we believe that the relationship $y(x)$ is in fact linear, then it would make more sense to find the "best" straight line that approximates the function. We need a criterion that allows us to determine the fitting function such that its deviation from the given points is minimized. The deviation Δ_i at point x_i is the difference between the fit $f(x)$ and the actual function $y(x)$, i.e.

$$\Delta_i = f(x_i) - y_i.$$

We examine a few optimization criteria, given n data points.

Minimization of $\sum_i \Delta_i$: Since errors may be both positive and negative, a positive error and negative error summed will give a sum that is less in magnitude than either error. Furthermore, since $-1+1=-2+2=-3+3=\cdots=0$ it is clear that the sum cannot give a unique minimum. This means that such a fitting criterion cannot allow the fitting function to be uniquely determined.

Minimization of $\sum_i |\Delta_i|$: When we have an error range (y_-, y_+), $y_- < y_+$ around the y coordinate of a data point (x, y), then the straight line fit $f(x) = mx + c$ passing between (x, y_-) and (x, y_+) such that $y_- \leq f(x) \leq y_+$ yields

$$|f(x)-y_-| + |f(x)-y_+| \leq |y_+ - f(x) + f(x) - y_-| = |y_+ - y_-|$$

which is independent of m and c. Again, the fitting function cannot be determined uniquely.

Minimization of $\sum_i \Delta_i^2$: Let $f(x)$ be a function fitting the data. Suppose we have two data points with the same x coordinates but with different y coordinates, namely (x, y_1) and (x, y_2). We have that

$$S = \sum_i \Delta_i^2 = (f(x)-y_2)^2 + (f(x)-y_1)^2 = e^2 + (e+d)^2,$$

where $e = f(x) - y_2$ and $d = y_2 + y_1$. Then it follows that

$$\frac{\mathrm{d}S}{\mathrm{d}e} = 2e + 2(e+d) = 2(2e+d)$$

and

$$\frac{\mathrm{d}^2S}{\mathrm{d}e^2} = 4 > 0.$$

Therefore S will be a *minimum* if $e = \frac{d}{2}$. This means that the value of $f(x)$ is the mean of the two y coordinates; a result which is both unique and intuitively acceptable and from mathematical statistics this norm is known to be the correct choice. The function $f(x)$ is therefore chosen so that

$$S = \sum_{i=0}^{n}[f(x_i) - y_i]^2 \tag{5.7}$$

is a minimum. We will see that, in the case of polynomial fitting, the norm (5.7) gives a unique result.

5.5 Least-squares polynomial fitting

A wide variety of functions may be used in least-squares curve fitting. We will leave the fitting of trigonometric, exponential and logarithmic functions etc. to the student as self-study, and discuss only the use of polynomials as fitting functions.

We consider the case where $f(x)$ in (5.7) is an mth degree polynomial $p_m(x)$, with $m \leq n$. Then $p_m(x)$ must be chosen such that

$$S = \sum_{i=0}^{n} [p_m(x_i) - y_i]^2 \tag{5.8}$$

is a *minimum*. Since

$$p_m(x) = a_0 + a_1 x + \cdots + a_m x^m \tag{5.9}$$

we must demand that

$$\frac{\partial S}{\partial a_k} = 0 \qquad \text{for } k = 0, 1, \ldots, m. \tag{5.10}$$

From (5.8) we then have

$$\frac{\partial S}{\partial a_k} = 2\sum_{i=0}^{n}[p_m(x_i) - y_i]\frac{\partial p_m}{\partial a_k}(x_i) \qquad \text{for } k = 0, 1, \ldots, m.$$

From (5.9) we have that

$$\frac{\partial p_m}{\partial a_k}(x_i) = x_i^k$$

and so

$$\frac{\partial S}{\partial a_k} = 2\sum_{i=0}^{n}[p_m(x_i) - y_i]x_i^k \qquad \text{for } k = 0, 1, \ldots, m. \tag{5.11}$$

Furthermore,

$$\frac{\partial^2 S}{\partial a_k^2} = 2\sum_{i=0}^{n} x_i^k x_i^k \geq 0 \qquad \text{for } k = 0, 1, \ldots, m,$$

so that the requirement in (5.10) does indeed give a minimum for S. This value is obtained from (5.10) and (5.11):

$$\sum_{i=0}^{n} p_m(x_i)x_i^k = \sum_{i=0}^{n} y_i x_i^k \qquad k = 0, 1, \ldots, m. \tag{5.12}$$

The system in (5.12) consists of $m+1$ equations in the $m+1$ unknowns a_k; the a_k are thus determined uniquely.

From (5.9) we obtain a more explicit form for (5.12):

$$\begin{array}{rcll}
a_0(n+1) + a_1\sum_i x_i + \cdots + a_m\sum_i x_i^m & = & \sum_i y_i & \quad (k=0) \\
a_0\sum_i x_i + a_1\sum_i x_i^2 + \cdots + a_m\sum_i x_i^{m+1} & = & \sum_i x_i y_i & \quad (k=1) \\
 & \vdots & & \\
a_0\sum_i x_i^m + a_1\sum_i x_i^{m+1} + \cdots + a_m\sum_i x_i^{2m} & = & \sum_i x_i^m y_i & \quad (k=m)
\end{array}$$

To fit a straight line, for example, we have $m = 1$ and

$$p_1(x) = a_0 + a_1 x$$

and the coefficients a_k are obtained from

$$\begin{aligned}
a_0(n+1) + a_1\sum_i x_i &= \sum_i y_i \\
a_0\sum_i x_i + a_1\sum_i x_i^2 &= \sum_i x_i y_i
\end{aligned} \tag{5.13}$$

Variance A question arises regarding the order of the fitting polynomial. If we increase the degree of the polynomial to n (number of data points $= n+1$) then the fitting polynomial becomes an interpolating polynomial, which has zero deviation (since an interpolating polynomial passes through each point). How may we measure the quality of the fit for various degrees? Mathematical statistics tells us that we choose the degree for which the variation σ^2 shows a minimum, where

$$\sigma^2 = \frac{\sum_i \Delta_i^2}{n-m}.$$

Example 5.2. Fit a straight line to the data points in the following table:

i	x_i	y_i
0	1	2.04
1	2	4.12
2	3	5.64
3	4	7.18
4	5	9.20
5	6	12.04

Using the data in the table, we find

$$\begin{aligned} n+1 &= 6 \\ \sum_i x_i &= 21 \\ \sum_i y_i &= 40.22 \\ \sum_i x_i^2 &= 91 \\ \sum_i x_i y_i &= 174.16 \end{aligned}$$

Substitution into equation (5.13) gives the two simultaneous equations

$$\begin{aligned} 6a_0 + 21a_1 &= 40.22 \\ 21a_0 + 91a_1 &= 174.16 \end{aligned}$$

which are solved to give

$$\begin{aligned} a_0 &= 0.0253 \\ a_1 &= 1.9080 \end{aligned}$$

The fit is thus

$$p_1(x) = 1.9080x + 0.0253.$$

Example 5.3. In the following table the coordinates of six data points are given. Fit polynomials of degrees 1 to 4 to this data, and decide which is the best fit.

x	0	0.8	1.4	2.1	2.7	3.4
y	0.015	0.644	1.926	4.442	7.274	11.621

The fitting polynomials, and variance for each, are given in the following table:

Degree	Fitting polynomial	σ^2
1	$p_1(x) = -1.601 + 3.416x$	2.1
2	$p_2(x) = 0.018 - 0.043x + 1.016x^2$	0.0010
3	$p_3(x) = 0.015 - 0.023x + 0.999x^2 + 0.003x^3$	0.0014
4	$p_4(x) = 0.017 - 0.113x + 1.142x^2 - 0.064x^3 + 0.010x^4$	0.0026

It is clear that the variance for all those with degree of two or higher are of the same order of magnitude, and that the optimal fit is the one with degree two.

5.6 Approximation with Chebyshev polynomials

One disadvantage of polynomial approximation is related to the fact that polynomial maxima and minima are spread unevenly on any interval: On $[-1, 1]$, $|x^k|$ has maxima at -1 and 1 and a minimum at 0. We now seek related functions that have evenly spaced maxima and minima, and for which the maxima and minima over a given interval are as small as possible. Possible candidates are, for example, the cosine functions: $\cos\theta, \cos 2\theta, \cos 3\theta, \ldots$.

5.6.1 Definition

The Chebyshev polynomial of the nth degree is defined by

$$T_n(x) = \cos(n \arccos x). \tag{5.14}$$

Here n is an integer and only the cases with $n \geq 0$ need to be studied, since it follows from the definition that $T_n = T_{-n}$.

We note that $T_n(x)$ is defined on $[-1, 1]$ only, due to the presence of the arccos function in (5.14). From the definition it follows

$$T_0(x) = 1 \qquad \text{and} \qquad T_1(x) = x.$$

Higher order Chebyshev polynomials are generated using a recursion formula, given by

$$T_{m+n}(x) + T_{m-n}(x) = 2T_m(x)T_n(x).$$

For the particular case $n = 1$ it follows that

$$T_{m+1}(x) = 2xT_m(x) - T_{m-1}(x) \tag{5.15}$$

and we find, for example,

$$\begin{aligned} T_2(x) &= 2xT_1(x) - T_0(x) \\ &= 2x^2 - 1 \\ T_3(x) &= 2xT_2(x) - T_1(x) \\ &= 4x^3 - 3x \end{aligned} \tag{5.16}$$

The relationships in (5.15) and (5.16) may be inverted to give powers of x in terms of T_n:

$$\begin{aligned} 1 &= T_0 \\ x &= T_1 \\ x^2 &= \frac{1}{2}(T_2 + 1) = \frac{T_2 + T_0}{2} \\ x^3 &= \frac{1}{4}(T_3 + 3x) = \frac{T_3 + 3T_1}{4} \end{aligned} \tag{5.17}$$

Zero points: From definition (5.14) it follows that

$$T_n(\xi) = 0 \qquad \text{for} \qquad n \arccos\ \xi = (2r+1)\frac{\pi}{2},$$

where r is an *integer*, and thus the zero points of $T_n(x)$ occur at

$$\xi_r = \cos\left(\frac{2r+1}{2n}\pi\right), \qquad r = 0, 1, \ldots, n-1. \tag{5.18}$$

We note that T_n has n zero points.

Extreme points: From definition (5.14) it follows that

$$T_n(x) = \pm 1 \qquad \text{for} \qquad n \arccos x = r\pi,$$

where r is an *integer* and thus we have the extreme points

$$x_r = \cos\left(\frac{r}{n}\pi\right), \qquad r = 0, 1, \ldots, n, \tag{5.19}$$

of T_n. We note that T_n has $n+1$ extreme points and $T_n(x_r) = (-1)^r$.

Orthogonality: Consider the integral

$$I_{mn} = \int_{-1}^{1} \frac{T_m(x)T_n(x)}{\sqrt{1-x^2}}\,\mathrm{d}x.$$

If we make the substitution

$$\begin{aligned} x &= \cos\theta \\ \mathrm{d}x &= -\sin\theta\ \mathrm{d}\theta \\ \sqrt{1-x^2} &= \sin\theta \end{aligned}$$

the integral becomes

$$\begin{aligned} I_{mn} &= -\int_{\pi}^{0} \frac{\cos m\theta \cos n\theta}{\sin\theta} \sin\theta\ \mathrm{d}\theta \\ &= \int_{0}^{\pi} \cos m\theta \cos n\theta\ \mathrm{d}\theta \\ &= \frac{1}{2}\int_{0}^{\pi} \left[\cos\big((m+n)\theta\big) + \cos\big((m-n)\theta\big)\right] \mathrm{d}\theta \end{aligned}$$

There are three distinct possibilities as far as the choice of m and n is concerned.

1. If $m = n = 0$, then
$$I_{mn} = \frac{1}{2}\int_{0}^{\pi} (1+1)\,\mathrm{d}\theta = \pi.$$
2. If $m = n \neq 0$, then
$$I_{mn} = \frac{1}{2}\int_{0}^{\pi} [\cos(2n\theta) + 1]\,\mathrm{d}\theta = \frac{\pi}{2}.$$
3. If $m \neq n$, then
$$I_{mn} = \frac{1}{2}\left[\frac{\sin(m+n)\theta}{m+n} + \frac{\sin(m-n)\theta}{m-n}\right]_{0}^{\pi} = 0.$$

In summary,

$$\int_{-1}^{1} \frac{T_m(x)T_n(x)}{\sqrt{1-x^2}}\,\mathrm{d}x = \begin{cases} \pi & m = n = 0 \\ \frac{\pi}{2} & m = n \neq 0 \\ 0 & m \neq n \end{cases} \tag{5.20}$$

Thus, the Chebyshev polynomials are said to be orthogonal with respect to the weight function $\frac{1}{\sqrt{1-x^2}}$ on the interval $[-1, 1]$.

5.6.2 Minimal property

From equations (5.16) it should be clear that the coefficient of x^n in $T_n(x)$ is equal to 2^{n-1}. The importance of Chebyshev polynomials arises from the following remarkable result:

Theorem 5.1. Consider the set $\mathcal{P}_n$ of all polynomials of degree n and with the coefficient of x^n equal to 1, on the interval $[-1, 1]$. If $p_n(x) \in \mathcal{P}_n$, then

$$\alpha_n = \max_{x\in[-1,1]} |p_n(x)|$$

is a minimum if

$$p_n(x) = \frac{1}{2^{n-1}} T_n(x).$$

Proof. Assume that there is a polynomial $p_n(x)$ with the coefficient of x^n equal to 1, that is $p_n(x) \in \mathcal{P}_n$, and for which

$$\alpha_n = \max_{x\in[-1,1]} |p_n(x)| < \left| \frac{T_n(x)}{2^{n-1}} \right| \le \frac{1}{2^{n-1}}$$

everywhere on $[-1, 1]$. The last inequality follows from definition (5.14) where $|T_n(x)| \le 1$ and so

$$\left| \frac{1}{2^{n-1}} T_n(x) \right| \le \frac{1}{2^{n-1}} \qquad \text{on } [-1, 1].$$

For the extreme points x_r, $r = 0, 1, \ldots, n$, from (5.19) we have

$$\begin{aligned} 2^{-n+1} T_n(x_0) - p_n(x_0) &> 0 \\ 2^{-n+1} T_n(x_1) - p_n(x_1) &< 0 \\ &\vdots \\ 2^{-n+1} T_n(x_n) - p_n(x_n) &> 0 \end{aligned}$$

for n even (the inequalities reverse for n odd). The function

$$F(x) = 2^{-n+1} T_n(x) - p_n(x)$$

thus changes sign n times on $[-1, 1]$ and so must have n roots. But, by construction, $F(x)$ is a polynomial of degree $n - 1$. Thus, the assumption regarding the existence of p_n is thus incorrect. Therefore, if $p_n(x) \in \mathcal{P}_n$, then

$$\alpha_n \ge \frac{T_n(x)}{2^{n-1}}$$

with equality only if α_n is a minimum. Q.E.D.

We may use theorem 5.1 to reduce the error in polynomial interpolation. If $p_n(x)$ is an interpolating polynomial that approximates a function $y(x)$ defined on $[-1, 1]$, recall that the error for polynomial interpolation is given by

$$y(x) - p_n(x) = \frac{y^{(n+1)}(\xi)}{(n+1)!} \prod_{i=0}^{n} (x - x_i).$$

It follows that

$$\max_{x\in[-1,1]} |y(x) - p_n(x)| \le \frac{1}{(n+1)!} \max_{x\in[-1,1]} \left| y^{(n+1)}(x) \right| \max_{x\in[-1,1]} \left| \prod_{i=0}^{n} (x - x_i) \right|.$$

Clearly, $\prod_{i=0}^{n}(x - x_i)$ is an element of $\mathcal{P}_n$ and therefore

$$\max_{x\in[-1,1]} \left| \prod_{j=0}^{n}(x - x_j) \right| \geq \frac{1}{2^{n-1}}.$$

One notices that the nodes x_i are the roots of the polynomial $\prod_{i=0}^{n}(x - x_j)$, and we may write

$$\frac{T_n(x)}{2^{n-1}} = \prod_{i=0}^{n}(x - \bar{x}_i)$$

where $\bar{x}_i$ are the zeroes of $T_n(x)$. We can thus better our interpolation if we replace the nodes x_i with the Chebyshev nodes $\bar{x}_i$.

We are not restricted to the interval $[-1, 1]$ and can apply this minimisation on an interval $[a, b]$ by the linear change of variable

$$x_j = \frac{\left[(a + b) + (b - a)\bar{x}_j\right]}{2}$$

for $j = 0, 1, \ldots, n - 1$.

Another application of theorem 5.1 is reducing the degree of a known approximation polynomial. Let $p_n(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + a_nx^n$ be a nth degree interpolating polynomial approximating the function $y(x)$. Consider now a $(n - 1)$th degree polynomial $p_{n-1}(x) = b_0 + b_1x + \cdots + b_{n-1}x^{n-1}$. We want $\max_{x\in[-1,1]} |p_n(x) - p_{n-1}(x)|$ to be as small as possible. Note that

$$\frac{p_n(x) - p_{n-1}(x)}{a_n} = \left(\frac{a_0 - b_0}{a_n}\right) + \left(\frac{a_1 - b_1}{a_n}\right) x + \cdots + \left(\frac{a_{n-1} - b_{n-1}}{a_n}\right) x^{n-1} + x^n$$

is indeed a monic polynomial. Since

$$\max_{x\in[-1,1]} \left| \frac{p_n(x) - p_{n-1}(x)}{a_n} \right| \geq \max_{x\in[-1,1]} \left| \frac{T_n(x)}{2^{n-1}} \right|$$

and the minimum will be obtained when we have equality, that is

$$\frac{p_n(x) - p_{n-1}(x)}{a_n} = \frac{T_n(x)}{2^{n-1}},$$

we can thus choose

$$p_{n-1}(x) = p_n(x) - \frac{a_n}{2^{n-1}}T_n(x).$$

5.6.3 Expansion of a function in terms of Chebyshev polynomials

We investigate the possibility of approximating a given function using an nth order expansion in terms of Chebyshev polynomials, i.e.

$$f(x) \approx \frac{1}{2}c_0T_0(x) + c_1T_1(x) + c_2T_2(x) + \cdots + c_nT_n(x). \tag{5.21}$$

There are two ways to obtain such an expansion, and we will illustrate each by means of an example.

Method 1: "Economization" of a power series. If the powers of x in a known series expansion are expressed in terms of Chebyshev polynomials and the series is truncated after the term $c_{n-1}T_{n-1}(x)$, then the resulting error will essentially be given by $c_nT_n(x)$ and so will oscillate between c_n and $-c_n$.

Example 5.4. Economize the Taylor series for the exponential function

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots .$$

We use (5.17) to write the powers of x in terms of T_n, i.e.

$$\begin{aligned} e^x = T_0 + T_1 + \frac{1}{2}\left(\frac{T_0 + T_2}{2}\right) + \frac{1}{6}\left(\frac{3T_1 + T_3}{4}\right) \\ + \frac{1}{24}\left(\frac{3T_0 + 4T_2 + T_4}{8}\right) + \frac{1}{120}\left(\frac{10T_1 + 5T_3 + T_5}{16}\right) + \cdots \end{aligned}$$

Next we collect terms in T_n for the first six terms, and truncate after T_3 to obtain

$$e^x \approx \frac{81}{64}T_0 + \frac{217}{192}T_1 + \frac{13}{48}T_2 + \frac{17}{384}T_3.$$

Finally use (5.16) to write the T_n in terms of powers of x:

$$e^x \approx \frac{191}{192} + \frac{383}{384}x + \frac{13}{24}x^2 + \frac{17}{96}x^3 = p(x).$$

If we truncate the Taylor series after the 3rd-order term, we obtain

$$e^x \approx \frac{192}{192} + \frac{384}{384}x + \frac{12}{24}x^2 + \frac{16}{96}x^3 = q(x).$$

It is clear that $p(x)$ and $q(x)$ are almost identical, but it is the subtle difference between the two that significantly affects the quality of the approximations. The error for these two approximations on $[-1, 1]$ is shown in the following table.

x	$\lvert e^x - p(x)\rvert$	$\lvert e^x - q(x)\rvert$
0	0.005	0
0.05	0.005	10^{-7}
0.1	0.005	10^{-6}
0.2	0.004	0.0001
0.5	0.002	0.003
0.8	0.005	0.02
1	0.007	0.05

It is clear that even though the Taylor series is better for small values of x, the Chebyshev series seems to "spread" the error evenly over the whole interval.

Method 2: Direct application of the orthogonality property. We multiply (5.21) by $\frac{T_k(x)}{\sqrt{1-x^2}}$ and integrate over $[-1, 1]$. From (5.20) it follows that

$$\int_{-1}^{1} \frac{f(x)T_k(x)}{\sqrt{1-x^2}}\,\mathrm{d}x = 0 + 0 + \cdots + \frac{\pi}{2}c_k + \cdots + 0$$

and so

$$c_k = \frac{2}{\pi}\int_{-1}^{1} \frac{f(x)T_k(x)}{\sqrt{1-x^2}}\,\mathrm{d}x.$$

The substitution $x = \cos\theta$ together with the definition (5.14) gives a more compact form for the integral

$$c_k = \frac{2}{\pi}\int_{0}^{\pi} f(\cos\theta)\cos(k\theta)\,\mathrm{d}\theta$$

An obvious disadvantage of this method is that sometimes difficult integrals may have to be determined.

Example 5.5. Expand $f(x) = \arcsin x$ in terms of Chebyshev polynomials.

We find the coefficients by

$$\begin{aligned} c_k &= \frac{2}{\pi}\int_0^{\pi} \arcsin(\cos\theta)\cos(k\theta)\,\mathrm{d}\theta \\ &= \frac{2}{\pi}\int_0^{\pi} \arcsin\left(\sin\left(\frac{\pi}{2}-\theta\right)\right)\cos(k\theta)\,\mathrm{d}\theta \\ &= \frac{2}{\pi}\int_0^{\pi} \left(\frac{\pi}{2}-\theta\right)\cos(k\theta)\,\mathrm{d}\theta \end{aligned}$$

We identify two cases:

1. If $k = 0$, then

$$c_0 = \frac{2}{\pi}\int_0^{\pi}\left(\frac{\pi}{2}-\theta\right)\,\mathrm{d}\theta = \frac{2}{\pi}\left[\frac{\pi}{2}\theta - \frac{\theta^2}{2}\right]_0^{\pi} = 0.$$

2. If $k \neq 0$, then

$$\begin{aligned} c_k &= \left[\frac{\sin(k\theta)}{k}\right]_0^{\pi} - \frac{2}{\pi}\left(\left[\frac{\theta\sin(k\theta)}{k}\right]_0^{\pi} - \int_0^{\pi}\frac{\sin(k\theta)}{k}\,\mathrm{d}\theta\right) \\ &= -\frac{2}{\pi}\left[\frac{\cos(k\theta)}{k^2}\right]_0^{\pi} = \frac{2}{\pi k^2}\left[1-(-1)^k\right]. \end{aligned}$$

It follows that $c_k = 0$ when k is even and $c_k = \frac{4}{\pi k^2}$ when k is odd.

We thus have

$$\begin{aligned} \arcsin x &= \frac{4}{\pi}\sum_{k=0}^{\infty}\frac{1}{(2k+1)^2}T_{2k+1}(x) \\ &= \frac{4}{\pi}\left[T_1(x) + \frac{T_3(x)}{9} + \frac{T_5(x)}{25} + \cdots\right]. \end{aligned}$$

If we truncated this series after j terms (i.e. when $k = j-1$), the last term in the series is

$$\frac{4T_{2j-1}(x)}{\pi(2j-1)^2}.$$

We know that $\max|T_{2j-1}| = 1$, so the residual term R of the series, i.e. the sum of all the terms for $k \geq j$, satisfies

$$\begin{aligned} |R| &= \frac{4}{\pi}\sum_{k=j}^{\infty}\frac{1}{(2k+1)^2}\,|T_{2k+1}(x)| \\ &\leq \frac{4}{\pi}\sum_{k=j}^{\infty}\frac{1}{(2k+1)^2} \\ &\leq \frac{4}{\pi}\left|\int_j^{\infty}\frac{1}{(2x+1)^2}\,\mathrm{d}x\right| \\ &= \frac{2}{(2j+1)\pi}. \end{aligned}$$

Hence, if a tolerance of ϵ was imposed, we would have

$$\begin{aligned} |R| &\leq \epsilon \\ \frac{2}{(2j+1)\pi} &\leq \epsilon \\ \therefore\ j &\geq \frac{1}{\pi\epsilon} - \frac{1}{2} \end{aligned}$$

and so we see that if ϵ is very small, then j must be very large for the desired tolerance to be satisfied.