LEAST SQUARES APPROXIMATION

Another approach to approximating a function $f(x)$ on an interval $a \le x \le b$ is to seek an approximation $p(x)$ with a small 'average error' over the interval of approximation. A convenient definition of the average error of the approximation is given by

$$E(p; f) \equiv \left[\frac{1}{b-a}\int_a^b [f(x) - p(x)]^2 \, dx\right]^{\frac{1}{2}} \qquad (1)$$

This is also called the *root-mean-square-error* (denoted subsequently by *RMSE*) in the approximation of $f(x)$ by $p(x)$. Note first that choosing $p(x)$ to minimize $E(p; f)$ is equivalent to minimizing

$$\int_a^b [f(x) - p(x)]^2 \, dx$$

thus dispensing with the square root and multiplying fraction (although the minimums are generally different). The minimizing of (1) is called the *least squares approximation problem*.

Example. Let $f(x) = e^x$, let $p(x) = \alpha_0 + \alpha_1 x$, α_0, α_1 unknown. Approximate $f(x)$ over $[-1, 1]$. Choose α_0, α_1 to minimize

$$g(\alpha_0, \alpha_1) \equiv \int_{-1}^{1} [e^x - \alpha_0 - \alpha_1 x]^2 \, dx \qquad (2)$$

$$g(\alpha_0, \alpha_1) = \int_{-1}^{1} \left\{ \begin{array}{c} e^{2x} + \alpha_0^2 + \alpha_1^2 x^2 - 2\alpha_0 e^x \\ -2\alpha_1 x e^x + 2\alpha_0 \alpha_1 x \end{array} \right\} dx$$

Integrating,

$$g(\alpha_0, \alpha_1) = c_1\alpha_0^2 + c_2\alpha_1^2 + c_3\alpha_0\alpha_1 + c_4\alpha_0 + c_5\alpha_1 + c_6$$

with constants $\{c_1, \ldots, c_6\}$, e.g.

$$c_1 = 2, \qquad c_6 = \left(e^1 - e^{-1}\right)/2.$$

g is a quadratic polynomial in the two variables α_0, α_1. To find its minimum, solve the system

$$\frac{\partial g}{\partial \alpha_0} = 0, \qquad \frac{\partial g}{\partial \alpha_1} = 0$$

It is simpler to return to (2) to differentiate, obtaining

$$2\int_{-1}^{1}\left[e^x-\alpha_0-\alpha_1 x\right](-1)\,dx=0$$
$$2\int_{-1}^{1}\left[e^x-\alpha_0-\alpha_1 x\right](-x)\,dx=0$$

This simplifies to

$$2\alpha_0=\int_{-1}^{1}e^x dx=e-e^{-1}$$
$$\frac{2}{3}\alpha_1=\int_{-1}^{1}xe^x dx=2e^{-1}$$

$$\alpha_0=\frac{e-e^{-1}}{2}\doteq 1.1752$$
$$\alpha_1=3e^{-1}\doteq 1.1036$$

Using these values for α_0 and α_1, we denote the resulting linear approximation by

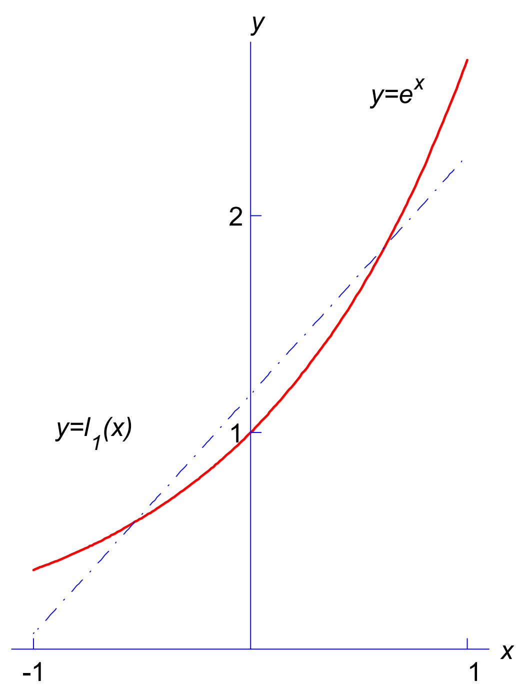
$$\ell_1(x)=\alpha_0+\alpha_1 x$$

It is called the best linear approximation to e^x in the *sense of least squares*. For the error,

$$\max_{-1\le x\le 1}\left|e^x-\ell_1(x)\right|\doteq 0.439$$

Errors in linear approximations of e^x:

Approximation	*Max Error*	*RMSE*
Taylor $t_1(x)$	0.718	0.246
Least squares $\ell_1(x)$	0.439	0.162
Chebyshev $c_1(x)$	0.372	0.184
Minimax $m_1(x)$	0.279	0.190



The linear least squares approximation to e^x

THE GENERAL CASE

Approximate $f(x)$ on $[a, b]$, and let $n \geq 0$. Seek $p(x)$ to minimize the *RMSE*. Write

$$p(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n$$

$$g(\alpha_0, \alpha_1, \ldots, \alpha_n) \equiv \int_{-1}^{1} \left[\begin{array}{r} f(x) - \alpha_0 - \alpha_1 x \\ - \cdots - \alpha_n x^n \end{array} \right]^2 dx$$

Find coefficients $\alpha_0, \alpha_1, \ldots, \alpha_n$ to minimize this integral. The integral $g(\alpha_0, \alpha_1, \ldots, \alpha_n)$ is a quadratic polynomial in the $n+1$ variables $\alpha_0, \alpha_1, \ldots, \alpha_n$.

To minimize $g(\alpha_0, \alpha_1, \ldots, \alpha_n)$, invoke the conditions

$$\frac{\partial g}{\partial \alpha_i} = 0, \qquad i = 0, 1, \ldots, n$$

This yields a set of $n+1$ equations that must be satisfied by a minimizing set $\alpha_0, \alpha_1, \ldots, \alpha_n$ for g. Manipulating this set of conditions leads to a simultaneous linear system.

To better understand the form of the linear system, consider the special case of $[a, b] = [0, 1]$. Differentiating g with respect to each α_i, we obtain

$$2\int_{-1}^{1} \left[e^x - \alpha_0 - \cdots - \alpha_n x^n\right](-1)\, dx = 0$$
$$2\int_{-1}^{1} \left[e^x - \alpha_0 - \cdots - \alpha_n x^n\right](-x)\, dx = 0$$
$$\vdots$$
$$2\int_{-1}^{1} \left[e^x - \alpha_0 - \cdots - \alpha_n x^n\right](-x^n)\, dx = 0$$

Then the linear system is

$$\sum_{j=0}^{n} \frac{\alpha_j}{i+j+1} = \int_0^1 x^i f(x)\, dx, \qquad i = 0, 1, \ldots, n$$

We will study the solution of simultaneous linear systems in Chapter 6. There we will see that this linear system is 'ill-conditioned' and is difficult to solve accurately, even for moderately sized values of n such as $n = 5$. As a consequence, this is not a good approach to solving for a minimizer of $g(\alpha_0, \alpha_1, \ldots, \alpha_n)$.

LEGENDRE POLYNOMIALS

Define the *Legendre polynomials* as follows.

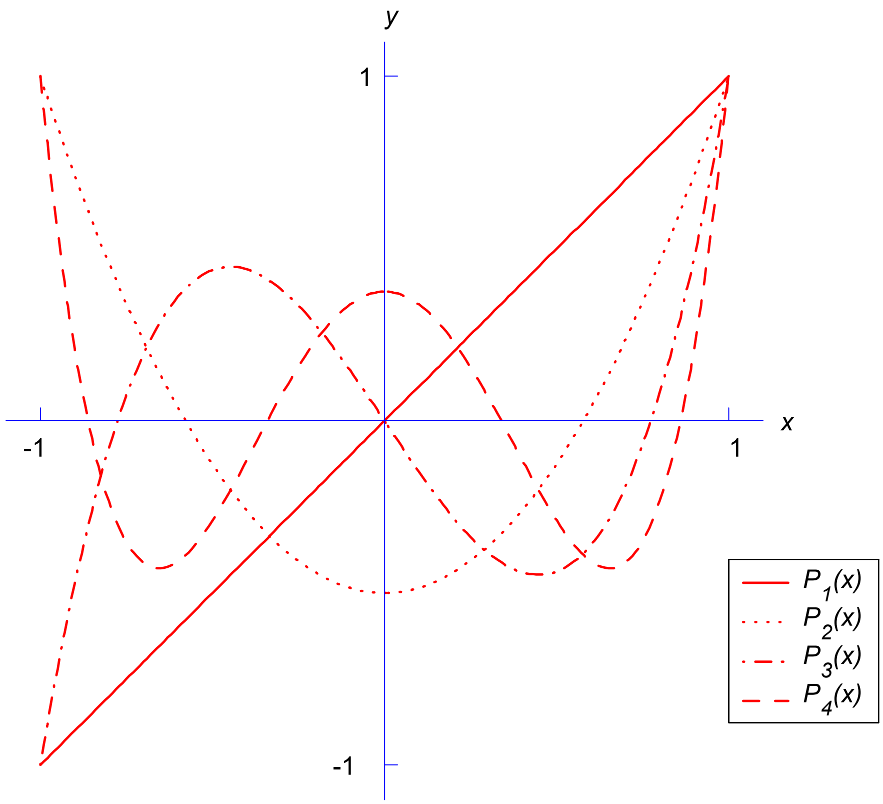
$$P_0(x) = 1$$

$$P_n(x) = \frac{1}{n!2^n} \cdot \frac{d^n}{dx^n}\left[\left(x^2 - 1\right)^n\right], \qquad n = 1, 2, \ldots$$

For example,

$$\begin{aligned} P_1(x) &= x \\ P_2(x) &= \frac{1}{2}\left(3x^2 - 1\right) \\ P_3(x) &= \frac{1}{2}\left(5x^3 - 3x\right) \\ P_4(x) &= \frac{1}{8}\left(35x^4 - 30x^2 + 3\right) \end{aligned}$$

The Legendre polynomials have many special properties, and they are widely used in numerical analysis and applied mathematics.



Legendre polynomials of degrees 1, 2, 3, 4

PROPERTIES

Introduce the special notation

$$(f, g) = \int_a^b f(x)g(x)\, dx$$

for general functions $f(x)$ and $g(x)$.

- *Degree and normalization*:

$$\deg P_n = n, \qquad P_n(1) = 1, \qquad n \geq 0$$

- *Triple recursion relation*: For $n \geq 1$,

$$P_{n+1}(x) = \frac{2n+1}{n+1} x P_n(x) - \frac{n}{n+1} P_{n-1}(x)$$

- *Orthogonality and size*:

$$\left(P_i, P_j\right) = \begin{cases} 0, & i \neq j \\ \dfrac{2}{2j+1}, & i = j \end{cases}$$

- *Zeroes*:

  All zeroes of $P_n(x)$ are located in $[-1, 1]$;
  all zeroes are simple roots of $P_n(x)$

- *Basis*: Every polynomial $p(x)$ of degree $\leq n$ can be written in the form

$$p(x) = \sum_{j=0}^{n} \beta_j P_j(x)$$

with the choice of $\beta_0, \beta_1, \ldots, \beta_n$ uniquely determined from $p(x)$:

$$\beta_j = \frac{\left(p, P_j\right)}{\left(P_j, P_j\right)}, \qquad j = 0, 1, \ldots, n$$

FINDING THE LEAST SQUARES APPROXIMATION

We solve the least squares approximation problem on only the interval $[-1, 1]$. Approximation problems on other intervals $[a, b]$ can be accomplished using a linear change of variable.

We seek to find a polynomial $p(x)$ of degree n that minimizes

$$\int_a^b [f(x) - p(x)]^2 \, dx$$

This is equivalent to minimizing

$$(f - p, f - p) \tag{3}$$

We begin by writing $p(x)$ in the form

$$p(x) = \sum_{j=0}^{n} \beta_j P_j(x)$$

$$p(x) = \sum_{j=0}^{n} \beta_j P_j(x)$$

Substitute into (3), obtaining

$$\begin{aligned} \widetilde{g}\,(\beta_0, \beta_1, \ldots, \beta_n) &\equiv (f - p, f - p) \\ &= \left(f - \sum_{j=0}^{n} \beta_j P_j, f - \sum_{i=0}^{n} \beta_i P_i \right) \end{aligned}$$

Expand this into the following:

$$\begin{aligned} \widetilde{g} &= (f, f) - \sum_{j=0}^{n} \frac{(f, P_j)^2}{(P_j, P_j)} \\ &+ \sum_{j=0}^{n} \left(P_j, P_j \right) \left[\beta_j - \frac{(f, P_j)}{(P_j, P_j)} \right]^2 \end{aligned}$$

Looking at this carefully, we see that it is smallest when

$$\beta_j = \frac{(f, P_j)}{(P_j, P_j)}, \qquad j = 0, 1, \ldots, n$$

Looking at this carefully, we see that it is smallest when

$$\beta_j = \frac{(f, P_j)}{(P_j, P_j)}, \qquad j = 0, 1, \ldots, n$$

The minimum for this choice of coefficients is

$$\widetilde{g} = (f, f) - \sum_{j=0}^{n} \frac{(f, P_j)^2}{(P_j, P_j)}$$

We call

$$\ell_n(x) = \sum_{j=0}^{n} \frac{(f, P_j)}{(P_j, P_j)} P_j(x) \tag{4}$$

the *least squares approximation of degree* n to $f(x)$ on $[-1, 1]$.

If $\beta_n = 0$, then its actual degree is less than n.

Example. Approximate $f(x) = e^x$ on $[-1, 1]$. We use (4) with $n = 3$:

$$\ell_3(x) = \sum_{j=0}^{3} \beta_j P_j(x), \quad \beta_j = \frac{(f, P_j)}{(P_j, P_j)} \tag{5}$$

The coefficients $\{\beta_0, \beta_1, \beta_2, \beta_3\}$ are as follows.

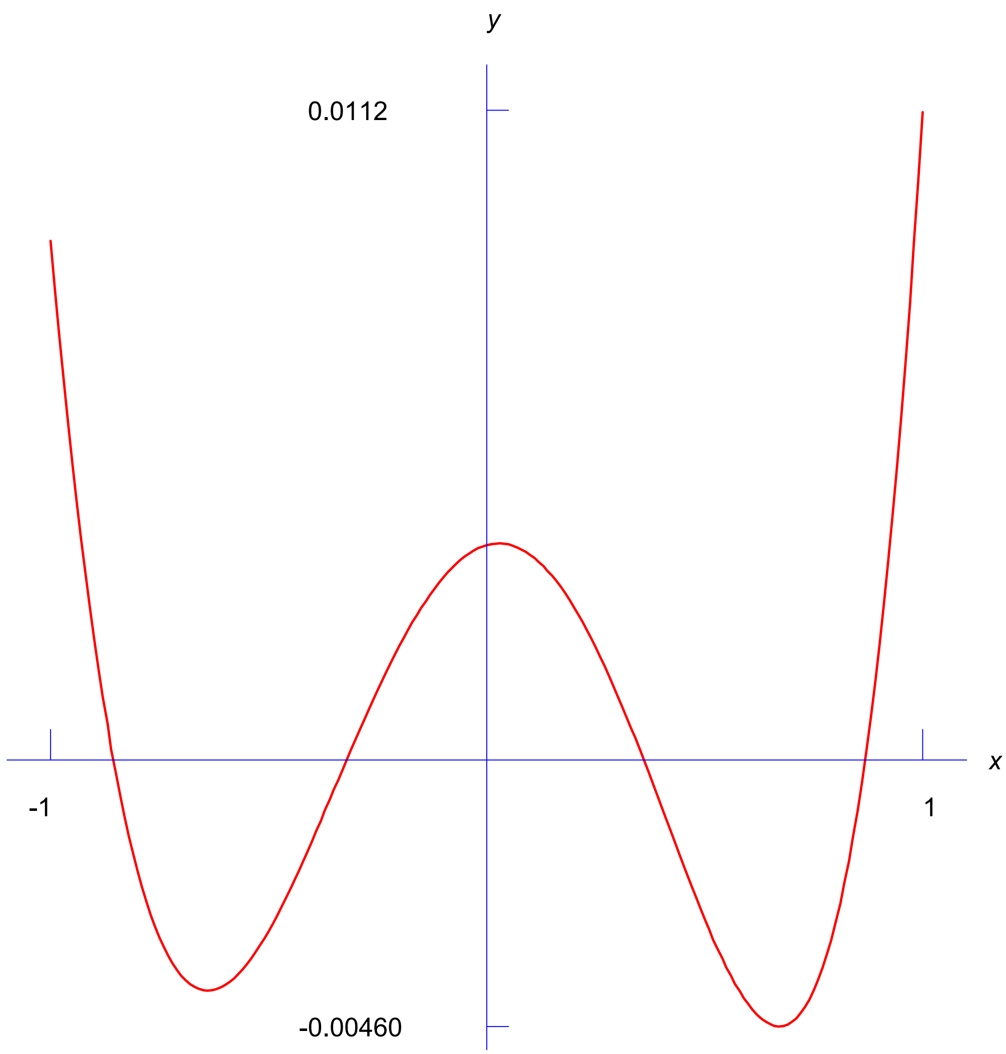
j	0	1	2	3
$\boldsymbol{\beta}_j$	2.35040	0.73576	0.14313	0.02013

Using (5) and simplifying,

$$\ell_3(x) = .996294 + .997955x + .536722x^2 + .176139x^3$$

The error in various cubic approximations:

Approximation	*Max Error*	*RMSE*
Taylor $t_3(x)$	.0516	.0145
Least squares $\ell_3(x)$	.0112	.00334
Chebyshev $c_3(x)$	.00666	.00384
Minimax $m_3(x)$	.00553	.00388



Error in the cubic least squares approximation to e^x