Matrices and linear equations

Elementary row operations used in Gaussian elimination

1. Add k times row i to row j (*notation*: $A_{ij}(k)$);
2. Multiply row i by k (*notation*: $M_i(k)$);
3. Switch rows i and j (*notation*: P_{ij})

Matrix algebra For addition, subtraction and multiplication of matrices the matrices must have a compatible size.

Matrix inverse The inverse of an $n \times n$ matrix, A, denoted A^{-1} satisfies $A^{-1}A = I_n = AA^{-1}$ where I_n is the $n \times n$ identity matrix with 1's along the diagonal and 0's everywhere else. To calculate A^{-1}, create an augmented matrix form by adding the identity matrix to the righthand side of the matrix A and perform Gaussian elimination (see below) until you get I_n on the lefthand side of the augmented matrix form.

Determinant The determinant of an $n \times n$ matrix A, denoted $|A|$, is a number which determines whether A^{-1} exists ($|A| \neq 0$) or not ($|A| = 0$). In fact there is a formula for A^{-1} in terms of the adjoint which is a matrix consisting of determinant. It is easy to calculate but long winded.

Non-singular matrices

The following four equivalent properties characterise a non-singular $n \times n$ matrix A:
(a) A has an inverse,
(b) the determinant of A is not zero,
(c) the linear system $A\boldsymbol{x} = \boldsymbol{b}$ has a unique solution for every column vector $\boldsymbol{b}$ of n elements,
(d) the linear system $A\boldsymbol{x} = \boldsymbol{0}$ has only the trivial solution $\boldsymbol{x} = \boldsymbol{0}$.

Gaussian elimination

Gaussian elimination without pivoting has the effect of expressing the coefficient matrix as $A = LU$, where L and U are, respectively, lower and upper triangular matrices; U is the coefficient matrix of the final triangular system, and L has 1 in each diagonal position and the multipliers used in the i-th stage of the elimination appear below the diagonal in the i-th column. To solve several sets of equations with the same coefficient matrix

A, but with different right-hand sides $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots$, we can calculate L and U once and then solve $L\boldsymbol{c}_1 = \boldsymbol{b}_1$ etc. by forward substitution, and $U\boldsymbol{x} = \boldsymbol{c}_1$ etc. by back substitution.

Examples (of Gaussian elimination)

1.

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 4 & 5 & 6 & 24 \\ 3 & 1 & -2 & 4 \end{array}\right) \overset{\substack{A_{12}(-4) \\ A_{13}(-3)}}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & -3 & -6 & -12 \\ 0 & -5 & -11 & -23 \end{array}\right) \overset{M_2(-\frac{1}{3})}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 1 & 2 & 4 \\ 0 & -5 & -11 & -23 \end{array}\right)$$

$$\overset{A_{23}(5)}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & -1 & -3 \end{array}\right) \overset{M_3(-1)}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \end{array}\right) \overset{\substack{A_{31}(-3) \\ A_{32}(-2)}}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

$$\overset{A_{21}(-2)}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

2. $$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 1 & 2 & 2 & 6 \\ 1 & 3 & 5 & 13 \end{array}\right) \overset{\substack{A_{12}(-1) \\ A_{13}(-1)}}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 0 & -1 & -3 \\ 0 & 1 & 2 & 4 \end{array}\right) \overset{P_{23}}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & -1 & -3 \end{array}\right)$$

3. $$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 4 & 5 & 6 & 24 \\ 2 & 7 & 12 & 40 \end{array}\right) \overset{\substack{A_{12}(-4) \\ A_{13}(-2)}}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & -3 & -6 & -12 \\ 0 & 3 & 6 & 22 \end{array}\right) \overset{M_2(-\frac{1}{3})}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 1 & 2 & 4 \\ 0 & 3 & 6 & 22 \end{array}\right)$$

$$\overset{A_{23}(-3)}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 0 & 10 \end{array}\right)$$

4. $$\left(\begin{array}{cccc|c} 1 & 3 & -5 & 1 & 4 \\ 2 & 5 & -2 & 4 & 6 \\ 1 & 1 & 11 & 4 & 3 \end{array}\right) \overset{\substack{A_{12}(-2) \\ A_{13}(-1)}}{\longrightarrow} \left(\begin{array}{cccc|c} 1 & 3 & -5 & 1 & 4 \\ 0 & -1 & 8 & 2 & -2 \\ 0 & -2 & 16 & 3 & -1 \end{array}\right) \overset{M_2(-1)}{\longrightarrow} \left(\begin{array}{cccc|c} 1 & 3 & -5 & 1 & 4 \\ 0 & 1 & -8 & -2 & 2 \\ 0 & -2 & 16 & 3 & -1 \end{array}\right)$$

$$\overset{A_{23}(2)}{\longrightarrow} \left(\begin{array}{cccc|c} 1 & 3 & -5 & 1 & 4 \\ 0 & 1 & -8 & -2 & 2 \\ 0 & 0 & 0 & -1 & 3 \end{array}\right) \overset{\substack{A_{21}(-3) \\ M_3(-1)}}{\longrightarrow} \left(\begin{array}{cccc|c} 1 & 0 & 19 & 7 & -2 \\ 0 & 1 & -8 & -2 & 2 \\ 0 & 0 & 0 & 1 & -3 \end{array}\right) \overset{\substack{A_{31}(-7) \\ A_{32}(2)}}{\longrightarrow} \left(\begin{array}{cccc|c} 1 & 0 & 19 & 0 & 19 \\ 0 & 1 & -8 & 0 & -4 \\ 0 & 0 & 0 & 1 & -3 \end{array}\right)$$

Partial pivoting

The LU decomposition described above is not always possible, even

when A is non-singular. Partial pivoting is used to avoid failure of Gaussian elimination through the occurrence of zero pivots. This also ensures that the multipliers do not exceed 1 in magnitude, which helps to reduce the effects of rounding errors. Gaussian elimination with partial pivoting has the effect of expressing A as LU, where U is again upper triangular, but now L is a permutation of a lower triangular matrix; this is always possible for a non-singular matrix.

Row scaling

Any equation in a linear system may be multiplied by a constant without affecting the solution, but the choice of pivots in Gaussian elimination may be affected. Row scaling is recommended, i.e., multiplying the equations by suitable constants to arrange that the element of largest magnitude in any row of the coefficient matrix is approximately 1.

Examples (of finding the inverse of a matrix)

1. $\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & -1 & 0 & 1 \end{array}\right) \xrightarrow{A_{12}(-3)} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -7 & -3 & 1 \end{array}\right)$

$$\xrightarrow{M_2(-\frac{1}{7})} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{3}{7} & -\frac{1}{7} \end{array}\right) \xrightarrow{A_{21}(-2)} \left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{7} & \frac{2}{7} \\ 0 & 1 & \frac{3}{7} & -\frac{1}{7} \end{array}\right)$$

2. $\left(\begin{array}{ccc|ccc} 1 & 1 & -1 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 & 0 \\ 0 & 0 & 3 & 0 & 0 & 1 \end{array}\right) \xrightarrow[M_2(\frac{1}{2})]{M_3(\frac{1}{3})} \left(\begin{array}{ccc|ccc} 1 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{3} \end{array}\right)$

$$\xrightarrow[A_{32}(-\frac{1}{2})]{A_{31}(1)} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & \frac{1}{3} \\ 0 & 1 & 0 & 0 & \frac{1}{2} & -\frac{1}{6} \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{3} \end{array}\right) \xrightarrow{A_{21}(-1)} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & 0 & \frac{1}{2} & -\frac{1}{6} \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{3} \end{array}\right)$$

3. $\left(\begin{array}{ccc|ccc} -2 & 1 & 1 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 & 1 & 0 \\ 1 & 1 & -2 & 0 & 0 & 1 \end{array}\right) \xrightarrow{M_1(-\frac{1}{2})} \left(\begin{array}{ccc|ccc} 1 & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 1 & -2 & 1 & 0 & 1 & 0 \\ 1 & 1 & -2 & 0 & 0 & 1 \end{array}\right)$

$$\xrightarrow[A_{13}(-1)]{A_{12}(-1)} \left(\begin{array}{ccc|ccc} 1 & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & -\frac{3}{2} & \frac{3}{2} & \frac{1}{2} & 1 & 0 \\ 0 & \frac{3}{2} & -\frac{3}{2} & \frac{1}{2} & 0 & 1 \end{array}\right) \xrightarrow{A_{23}(1)} \left(\begin{array}{ccc|ccc} 1 & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & -\frac{3}{2} & \frac{3}{2} & -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{array}\right)$$

The rule for calculating the determinant

Take any row (or column) and working from left to right (or top to bottom) successively take each number and multiply it by the smaller determinant you get by deleting the row and column (or column and row) which contain the number. Finally take the alternating sum of the smaller determinants using the appropriate signs from the following matrix

$$\begin{pmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Elementary row operations can be performed on determinants with the following change to $|A|$ 1. no change; 2. $|A|$ is multiplied by k; 3. $|A|$ is multiplied by -1. With the word "row" replaced by "column" all of the previous properties hold.

Examples (of calculating the determinant of a matrix)

1. $$\begin{vmatrix} {}^{+}2 & {}^{-}-1 & {}^{+}3 \\ 4 & 1 & 0 \\ 3 & 4 & 1 \end{vmatrix} = 2\underbrace{\begin{vmatrix} 1 & 0 \\ 4 & 1 \end{vmatrix}}_{=1} - (-1)\underbrace{\begin{vmatrix} 4 & 0 \\ 3 & 1 \end{vmatrix}}_{=4} + 3\underbrace{\begin{vmatrix} 4 & 1 \\ 3 & 4 \end{vmatrix}}_{=16-3} = 45$$

2. $$\begin{vmatrix} {}^{+}-2 & {}^{-}1 & {}^{+}1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{vmatrix} = -2\underbrace{\begin{vmatrix} -2 & 1 \\ 1 & -2 \end{vmatrix}}_{=4-1} - 1\underbrace{\begin{vmatrix} 1 & 1 \\ 1 & -2 \end{vmatrix}}_{=-2-1} + 1\underbrace{\begin{vmatrix} 1 & -2 \\ 1 & 1 \end{vmatrix}}_{1+2} = 0$$

3. $$\begin{vmatrix} {}^{+}2 & {}^{-}-1 & {}^{+}3 \\ {}^{-}4 & {}^{+}1 & {}^{-}0 \\ {}^{+}3 & {}^{-}4 & {}^{+}1 \end{vmatrix} = -4\underbrace{\begin{vmatrix} -1 & 3 \\ 4 & 1 \end{vmatrix}}_{=-1-12} + \underbrace{\begin{vmatrix} 2 & 3 \\ 3 & 1 \end{vmatrix}}_{=2-9} = 45$$

   or $$= 3\underbrace{\begin{vmatrix} 4 & 1 \\ 3 & 4 \end{vmatrix}}_{=16-3} + 1 \times \underbrace{\begin{vmatrix} 2 & -1 \\ 4 & 1 \end{vmatrix}}_{=2+4} = 45$$

Hints and tips (for determinants)

Use Gaussian elimination to create well-placed zeros; If all but one number in a row (or column) is zero, then expansion about that row (or column) reduces the size of the problem by 1; If one row (or column) is a multiple of another row (or column) then the determinant is 0; If the matrix is completely 0 below (or above) the main diagonal, then $|A|$ is the product of the diagonal numbers.

Eigenvalues and eigenvectors

Eigenvalues and Eigenvectors Let A be an $n \times n$ matrix, the eigenvalues of A, $\{\lambda_i\}_{i=1}^n$, satisfy $|A - \lambda I| = 0$; the polynomial $\det(A - \lambda I)$ is called the characteristic polynomial of the matrix A. The corresponding eigenvectors, $\{\boldsymbol{w}^i\}_{i=1}^n$, may be found by finding the non-zero solution of $(A - \lambda_i I)\boldsymbol{w}^i = \boldsymbol{0}$ (the eigenvector should include an unknown b_i). Geometrically eigenvectors are special vectors which when pre-multiplied by A preserve their direction and get scaled up/down in accordance with the eigenvalue.

Hints and Tips 1. If your eigenvector *is* zero (i.e. you don't have a b_i in your answer) then you have either (a) *not* found the right eigenvalues or (b) made an error in Gaussian elimination; 2. If A consists of real numbers and $\{\lambda, \boldsymbol{w}\}$ is an eigenvalues/eigenvectors then so is $\{\overline{\lambda}, \overline{\boldsymbol{w}}\}$; 3. Moreover, if $A^T = A$ (symmetric) then all of the eigenvalues are real; 4. $|A| = \lambda_1 \cdots \lambda_n$; 5. A does not necessarily have n eigenvectors.

Theorem Let the eigenvectors of A be $\{\boldsymbol{w}^i\}_{i=1}^n$, $W = (\boldsymbol{w}^1 \cdots \boldsymbol{w}^n)$ and let the eigenvectors form a basis ($|W| \neq 0$) then

1. The solution to $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{z} = A\boldsymbol{z}$ is $\boldsymbol{z}(t) = e^{\lambda_1 t}\boldsymbol{w}^1 + \cdots + e^{\lambda_n t}\boldsymbol{w}^n$;
2. $W^{-1}AW$ is a diagonal matrix with the eigenvalues on the diagonal.

Moreover, if A is symmetric or $\{\lambda_i\}_{i=1}^n$ are all different then the eigenvectors form a basis.

Method for solving $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{z} = A\boldsymbol{z}$, $\boldsymbol{z}(0) = \boldsymbol{a}$

1. Find the eigenvalues of A, i.e. solve $|A - \lambda I| = 0$;
2. Find the eigenvectors corresponding to the eigenvalues;
3. $\boldsymbol{z}(t) = e^{\lambda_1 t}\boldsymbol{w}^1 + \cdots + e^{\lambda_n t}\boldsymbol{w}^n$;
4. Set $t = 0$ and solve $\boldsymbol{z}(0) = \boldsymbol{a}$ for $b_1, \cdots, b_n$ (b_i is the unknown coefficient in $\boldsymbol{w}^i$).

The Cayley-Hamilton theorem states that every square matrix satisfies its characteristic equation.

Iterative methods for linear equations

For large sparse systems of equations of the form $A\boldsymbol{x} = \boldsymbol{b}$, iterative methods are useful. Only the two simplest methods are mentioned here. The matrix A may be written as $D + L + U$ where D is diagonal, L is lower triangular and U is upper triangular.

Jacobi's method $$D\boldsymbol{x}^{(k+1)} = \boldsymbol{b} - (L+U)\boldsymbol{x}^{(k)}.$$

The Gauss-Seidel method $$(D+L)\boldsymbol{x}^{(k+1)} = \boldsymbol{b} - U\boldsymbol{x}^{(k)}.$$

A convergence theorem

The Jacobi and Gauss-Seidel iterates converge to the solution of a set of n linear equations in n unknowns if the $n \times n$ coefficient matrix is diagonally dominant.

An example

$$\begin{aligned} 10x_1 + x_2 + x_3 &= 12 \\ x_1 + 10x_2 - x_3 &= 10 \\ -x_1 - x_2 + 10x_3 &= 8 \end{aligned}$$

Jacobi's method:

$$\begin{aligned} 10x_1^{(k+1)} &= 12 - x_2^{(k)} - x_3^{(k)} \\ 10x_2^{(k+1)} &= 10 - x_1^{(k)} + x_3^{(k)} \\ 10x_3^{(k+1)} &= 8 + x_1^{(k)} + x_2^{(k)} \end{aligned}$$

$\boldsymbol{x}^{(0)}$	$\boldsymbol{x}^{(1)}$	$\boldsymbol{x}^{(2)}$	$\boldsymbol{x}^{(3)}$	$\boldsymbol{x}^{(4)}$	$\boldsymbol{x}^{(5)}$
0	1.2	1.02	1.002	1.0002	1.00002
0	1.0	0.96	1.000	0.9996	1.00000
0	0.8	1.02	0.998	1.0002	0.99998

Gauss-Seidel method:

$$\begin{aligned} 10x_1^{(k+1)} &= 12 - x_2^{(k)} - x_3^{(k)} \\ 10x_2^{(k+1)} &= 10 - x_1^{(k+1)} + x_3^{(k)} \\ 10x_3^{(k+1)} &= 8 + x_1^{(k+1)} + x_2^{(k+1)} \end{aligned}$$

$\boldsymbol{x}^{(0)}$	$\boldsymbol{x}^{(1)}$	$\boldsymbol{x}^{(2)}$	$\boldsymbol{x}^{(3)}$	$\boldsymbol{x}^{(4)}$
0	1.2	1.0112	0.99992	0.99999
0	0.88	0.99968	1.00012	1.000001
0	1.008	1.00109	1.000004	0.999999

Spline functions

A spline of first degree is a piecewise linear function and a quadratic spline consists of a piecewise quadratic function which has a continuous first derivative at the knots. A cubic spline is a piecewise cubic function which has continuous first and second derivatives at the knots. An interpolating spline takes prescribed values at the knots.

A cubic spline with knots x_0, x_1, ..., x_n, where $a = x_0 < x_1 < x_2 < \cdots < x_n = b$, is a function $S(x)$ having the following properties:

(i) on each subinterval $I_j = [x_j, x_{j+1}]$, for $j = 0, 1, \ldots, n-1$, $S(x)$ is a cubic polynomial,

(ii) $S(x)$ is twice continuously differentiable for $a < x < b$.

An interpolating cubic spline, agreeing with some given function $f(x)$ at the knots, also satisfies the interpolation conditions

(iii) $S(x_j) = f(x_j)$ for $j = 0, 1, \ldots, n$.

Let $f_j = f(x_j)$, $h_j = x_{j+1} - x_j$ and $k_j = S'(x_j)$ for the relevant values of j; the derivatives k_j at the knots have to be computed as part of the calculation. On the interval $[x_j, x_{j+1}]$, the spline function is a cubic polynomial which may be written as

$$S_j(x) = f_j + k_j(x - x_j) + a_{j2}(x - x_j)^2 + a_{j3}(x - x_j)^3.$$

To satisfy the conditions $S(x_j) = f_j$, $S(x_{j+1}) = f_{j+1}$, $S'(x_j) = k_j$ and $S'(x_{j+1}) = k_{j+1}$, we must have

$$a_{j2} = \frac{3(f_{j+1} - f_j)}{h_j^2} - \frac{2k_j + k_{j+1}}{h_j}, \qquad a_{j3} = \frac{2(f_j - f_{j+1})}{h_j^3} + \frac{k_j + k_{j+1}}{h_j^2}.$$

Continuity of the second derivative at the interior knots imposes $n - 1$ conditions, in the form of the equations

$$\frac{k_{j-1}}{h_{j-1}} + 2\left(\frac{1}{h_{j-1}} + \frac{1}{h_j}\right)k_j + \frac{k_{j+1}}{h_j} = \frac{3(f_{j+1} - f_j)}{h_j^2} + \frac{3(f_j - f_{j-1})}{h_{j-1}^2}, \text{ for } j = 1, 2, \ldots, n-1.$$

This is a set of $n-1$ equations (a tridiagonal system) for the $n+1$ derivative values k_j, for $j = 0, 1, \ldots, n$; consequently two of those values must be specified in some other way.

Possible end conditions are:

1. We provide numerical values for k_0 and k_n; possibly $k_0 = f'(x_0)$ and $k_n = f'(x_n)$, if known.

2. The so-called "natural cubic spline" has $S''(x_0) = 0 = S''(x_n)$. That requires

$$2k_0 + k_1 = 3(f_1 - f_0)/h_0 \qquad \text{and} \qquad k_{n-1} + 2k_n = 3(f_n - f_{n-1})/h_n.$$

3. The "not a knot" condition requires that the *third* derivative, $S'''(x)$, be continuous at x_1 and x_{n-1}, which means that the first and last interior knots are not active. This is the method used by the MATLAB m-file spline, called by choosing the spline option in `interp1`.

For evenly spaced knots, with the constant knot spacing h, the continuity conditions simplify to

$$k_{j-1} + 4k_j + k_{j+1} = 3(f_{j+1} - f_{j-1})/h, \text{ for } j = 1, 2, \ldots, n-1.$$

An example
Here we consider cubic spline approximation for $\sin x$ on the interval $[0, \pi]$, with the knots $x_0 = 0$, $x_1 = \pi/4$, $x_2 = \pi/2$, $x_3 = 3\pi/4$ and $x_4 = \pi$. The continuity equations are obtained by putting $h = \pi/4$ in the equation just above, and using $f_0 = 0$, $f_1 = 1/\sqrt{2}$, $f_2 = 1$, $f_3 = 1/\sqrt{2}$ and $f_4 = 0$. They are

$$\begin{aligned} k_0 + 4k_1 + k_2 &= \frac{12}{\pi}(f_2 - f_0) = \frac{12}{\pi}, \\ k_1 + 4k_2 + k_3 &= \frac{12}{\pi}(f_3 - f_1) = 0, \\ k_2 + 4k_3 + k_4 &= \frac{12}{\pi}(f_4 - f_2) = -\frac{12}{\pi}. \end{aligned}$$

If we take the end conditions $k_0 = f'(x_0) = 1$ and $k_4 = f'(x_4) = -1$, the equations become

$$4k_1 + k_2 = 12/\pi - 1, \qquad k_1 + 4k_2 + k_3 = 0, \qquad k_2 + 4k_3 = 1 - 12/\pi.$$

From these we get $k_2 = 0$ and $k_1 = -k_3 = 3/\pi - 1/4 = 0.704930$, correct to 6 decimal places.

On $[0, \pi/4]$ the spline function is

$$S_0(x) = f_0 + k_0 x + a_{02}x^2 + a_{03}x^3,$$

$$\text{with} \quad a_{02} = \frac{3(f_1 - f_0)}{h^2} - \frac{2k_0 + k_1}{h} = -0.005068, \qquad a_{03} = -0.155147,$$

both correct to six decimal places. Therefore

$$S_0(x) = x - 0.005068x^2 - 0.155147x^3.$$

See §18.4 Kreyszig p. 952 for other worked examples.

Errors

As a particular case of Taylor's theorem (the Mean Value Theorem), if a number a is approximated by another number A — e.g., π approximated by 3.14159 — the resulting change in the value of a differentiable function f is

$$f(a) - f(A) = (a - A)f'(c), \qquad \text{for some } c \text{ between } a \text{ and } A.$$

Ill-conditioned problems

A problem is said to be *ill-conditioned* if small changes in the data associated with the problem can cause relatively large changes in the solution. A condition number of a matrix A is a measure of the sensitivity of the solution of a set of linear equations $A\boldsymbol{x} = \boldsymbol{b}$ to changes in the data, and consequently to any rounding or other errors arising while solving the equations.

Loss of significance

This results from subtraction of numbers which are nearly equal and are not known exactly. The error from this source may be reduced by rearrangement, e.g., $\sqrt{a} - \sqrt{b} = (a - b)/(\sqrt{a} + \sqrt{b})$, or by expansion, e.g.,

$$\frac{2(1 - \cos x) - x^2}{x^4} = -\frac{1}{12} + \frac{x^2}{360} - \frac{x^4}{20160} + \cdots$$

Infinite integrals and singular integrands

Integrals over infinite intervals may be approximated by specially designed formulae, such as the Gauss-Laguerre formulae. Alternatively, we might change the variable of integration so as to convert the integration interval to a finite one, or else we could use one of the standard integration formulae on a finite part of the interval and combine this with a bound on the remainder, showing that its neglect is consistent with the desired accuracy. Newton-Cotes methods can't be applied directly to an integrand which is undefined at an end-point of the interval, but a suitable change of variable may remove the singularity, or we may be able to subtract off the singular part and integrate it exactly.

Numerical differentiation

Some simple formulae for the first derivative, with their truncation errors, are

$$f'(x) \approx \frac{f(x+h)-f(x)}{h}, \qquad \text{truncation error } -\frac{h}{2}f''(\zeta), \qquad x<\zeta<x+h$$

$$f'(x) \approx \frac{f(x)-f(x-h)}{h}, \qquad \text{truncation error } \frac{h}{2}f''(\theta), \qquad x-h<\theta<x$$

$$f'(x) \approx \frac{f(x+h)-f(x-h)}{2h}, \qquad \text{truncation error } -\frac{h^2}{6}f'''(\eta), \qquad x-h<\eta<x+h$$

The third one is a second-order formula and the others are of first order.

Derivatives of higher order may similarly be approximated. For example (see one of the problems)

$$f''(x) \approx \frac{f(x+h)-2f(x)+f(x-h)}{h^2}.$$

Rounding errors cause difficulty in estimating derivatives numerically. The formulae involve subtraction of function values which not known exactly and which, for sufficiently small steplengths h, are nearly equal. It is therefore important to avoid steplengths that are too small. Richardson extrapolation can help in this respect.