Chapter 2

Mathematical Background

Core Topics

(i) Concepts from Calculus (2.2)

(ii) Vectors (2.3)

(iii) Matrices and Linear Algebra (2.4)

(iv) Ordinary Differential Equations (ODEs) (2.5)

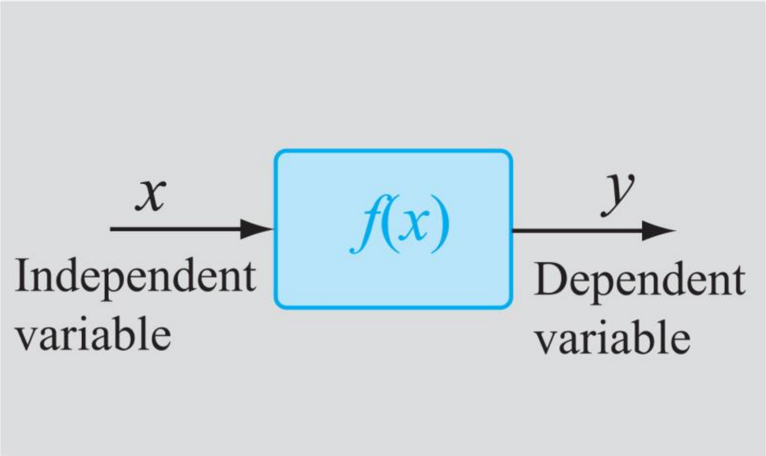
(v) Functions of two or more independent variables and partial differentation (2.6)

(vi) Taylor Series (2.7)

(vii) Inner Product and Orthogonality (2.8)

Concepts from Calculus

Functions:



- Consider $y = f(x)$. y is the dependent variable and x is the independent variable.
- A function may have many independent variables. Consider $g(x, y, z)$.
- The span of values that x and y can have are known as the domain (of x) and range (of y) respectively.
- Also called intervals: the open interval $]a, b[$ (or sometimes (a, b)) and the closed interval $[a, b]$.

Limit of a Function:

Function without any breaks

For continuous functions:

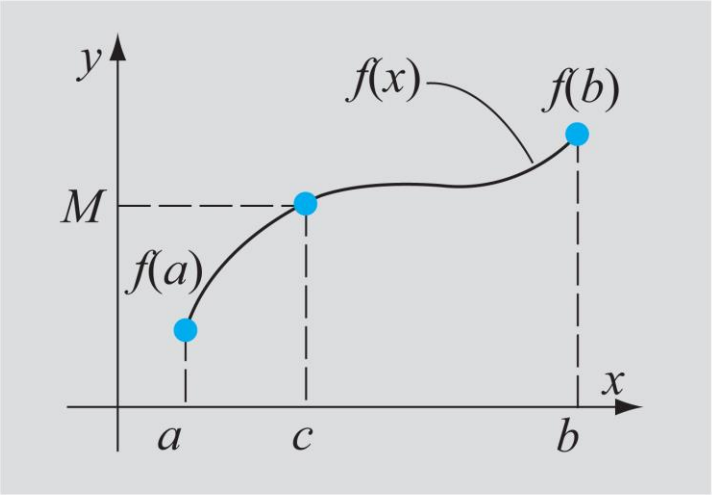
$$\lim_{x \to a} f(x) = f(a) = L$$

- Note $f(x)$ may be discontinuous or singular at a. In these cases the limit is said not to exist.

Continuity of a Function:

- A function $f(x)$ at $x = a$ is said to be continuous if
  - $f(a)$ exists
  - $\lim_{x \to a} f(x) = f(a)$

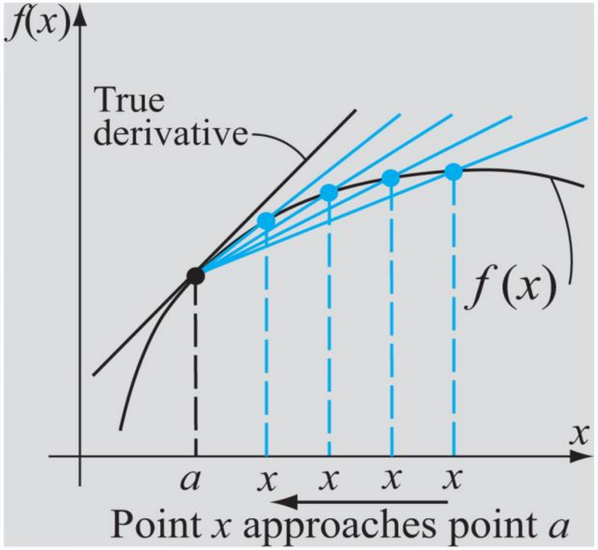
The Intermediate Value Theorem:



- States that if $f(x)$ is continuous in $[a, b]$ and M is any number between $f(a)$ and $f(b)$ then $\exists\, c \in [a, b]$ such that $f(c) = M$.

there exists

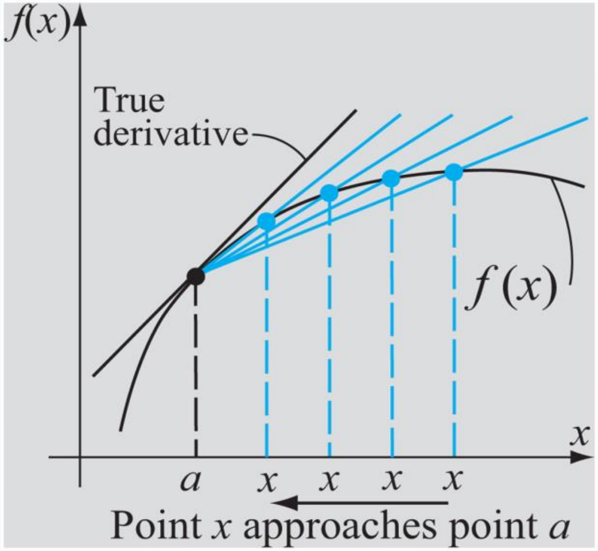
Derivatives of a Function:



- The derivative of a function $y = f(x)$ is denoted $\frac{dy}{dx}$, $\frac{df(x)}{dx}$ or $f'(x)$ and is defined as:

$$\left.\frac{dy}{dx}\right|_{x=a} = f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

Derivatives of a Function:



- $\frac{f(x)-f(a)}{x-a}$ - slope of the secant connecting $(a, f(a))$ and $(x, f(x))$
- In the limit as $x \to a$ we get slope of tangent at $(a, f(a))$
- This slope is equivalent to the rate of change of $f(x)$ w.r.t. x at $x = a$

Derivatives of a Function:

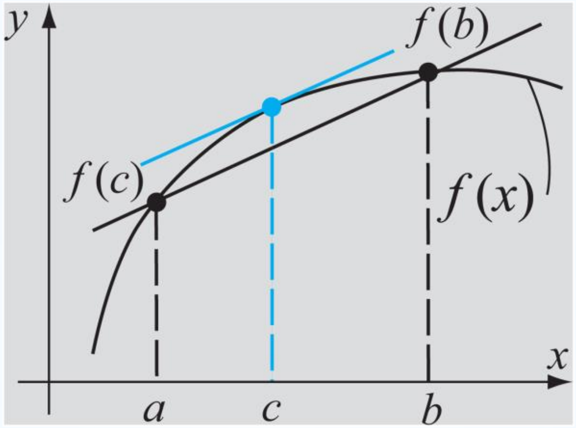
- ***The Chain Rule:*** if $y = f(u)$ and $u = g(x)$ then:

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$$

- ***The Product Rule:***

$$\frac{du(x)v(x)}{dx} = u(x)\frac{dv(x)}{dx} + v(x)\frac{du(x)}{dx}$$

The Mean Value Theorem (for Derivatives):



States that if $f(x)$ is a continuous function in $[a, b]$ and is differentiable in the open interval $]a, b[$ then $\exists$ c $\in]a, b[$ such that:

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Integral (Antiderivative) of a Function:

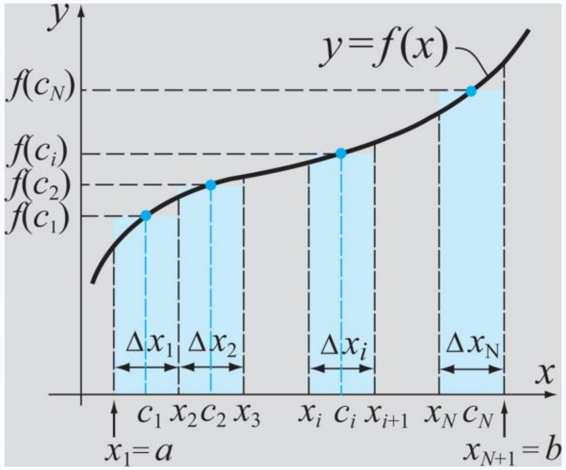
- If $f(x) = \frac{dF(x)}{dx}$ then the ***antiderivative*** or ***indefinite integral*** of $f(x)$ is $\mathrm{F}(x)$ and is written:

$$F(x) = \int f(x)dx$$

- The ***definite integral*** $'\boldsymbol{I}'$ is denoted by:

$$I = \int_a^b f(x)dx$$

Integral (Antiderivative) of a Function:



Assuming $f(x)$ is defined and continuous over $[a, b]$ then:

$$\int_a^b f(x)dx = \lim_{\Delta x_i \to 0} \sum_{i=1}^{n} f(c_i)\Delta x_i$$

and

$$\int f(x)dx = \int_{-\infty}^{+\infty} f(x)dx = \lim_{\Delta x_i \to 0} \sum_{i=-\infty}^{\infty} f(c_i)\Delta x_i$$

$$\sum_{i=1}^{n} f(c_i)\Delta x_i \; - Riemann\ Sum$$

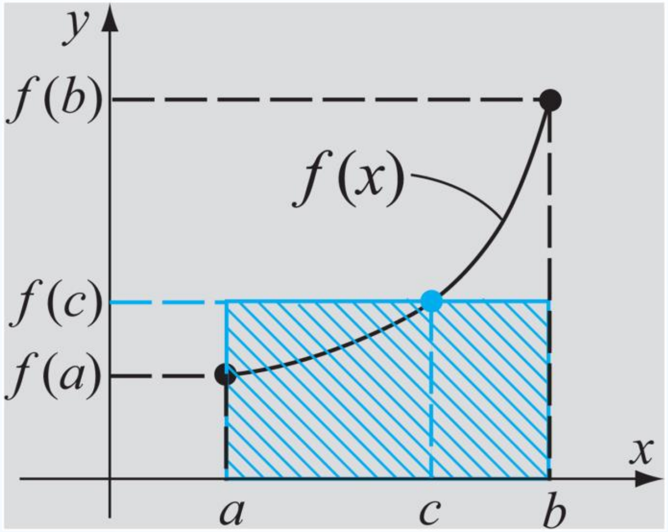
Fundamental Theorems of Calculus:

$$\int_a^b f(x)dx = F(b) - F(a)$$

and

$$\frac{d}{dx}\left(\int_a^x f(\xi)d\xi\right) = f(x)$$

Mean Value Theorem for Integrals:



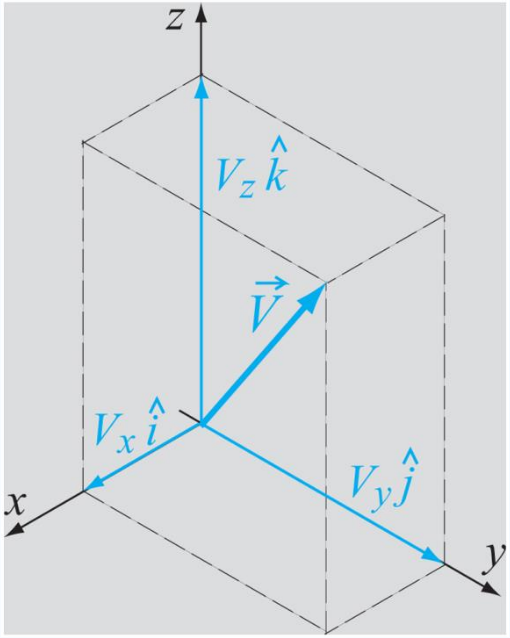
If $f(x)$ is continuous over $[a, b]$ then $\exists\, c \in [a, b]$ such that:

$$\int_a^b f(x)dx = f(c)(b - a)$$

The average value of $f(x)$ over the interval is given by:

$$\langle f \rangle = \frac{1}{b - a} \int_a^b f(x)dx$$

Vectors



- ***Vectors*** are quantities that have a ***magnitude*** and a ***direction***.
  Examples are force, velocity and acceleration.

- ***Scalars*** are quantities that have a ***magnitude*** only.
  Examples are mass, speed and height.

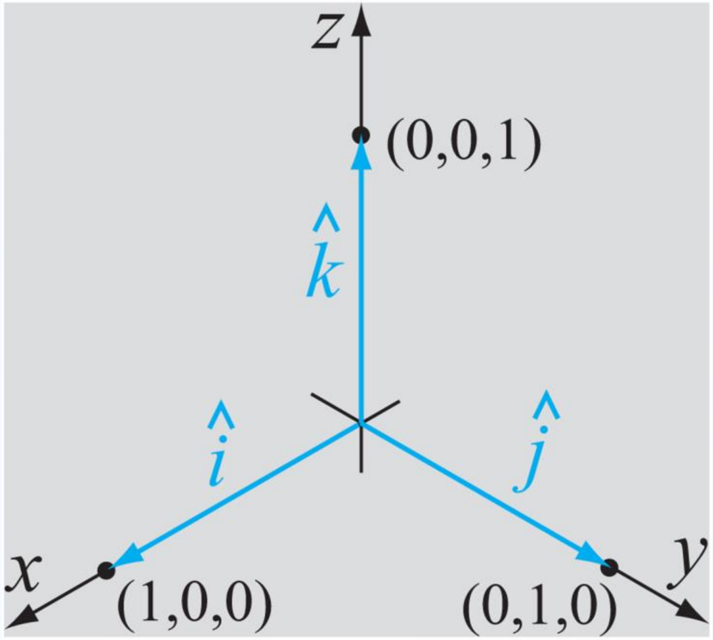
Vectors:

- To denote the vector 'V' we write $\vec{V}$. The magnitude of the vector is denoted V or $|V|$

- Once a coordinate system has been chosen, a vector may be represented graphically as a directed line segment (as in the diagram in the previous slide)

- Projections of the vector onto each of the coordinate axes define the ***components*** of the vector.

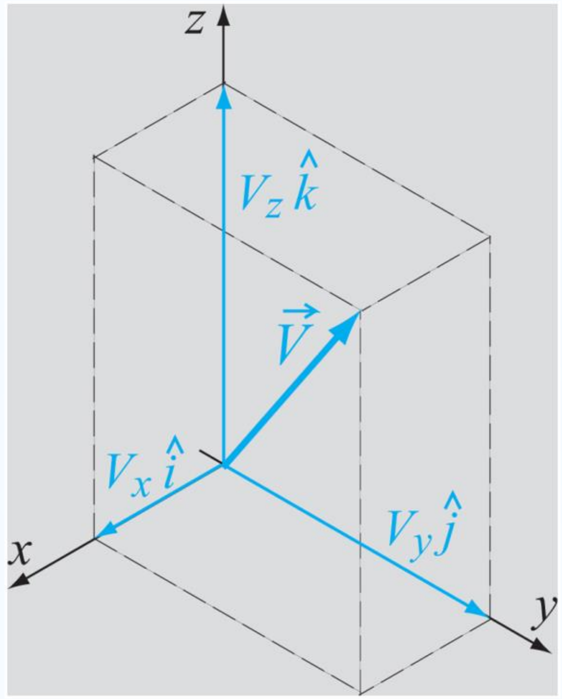
Vectors:



- $\hat{\imath}, \hat{\jmath}$, and $\hat{k}$ are the unit vectors in the x, y and z directions respectively

- A unit vector is a vector with a magnitude of unity in a particular direction e.g. $\hat{v} = \frac{\vec{V}}{|V|}$

Vectors:



We may then write:

$$\vec{V} = V_x\hat{\imath} + V_y\hat{\jmath} + V_Z\hat{k}$$

where $V_x\hat{\imath}$ is the x-component of $\vec{V}$ (the projection of $\vec{V}$ onto the x-axis) and so on.

Vectors:

A vector may be written by listing the magnitudes of its components in a row or column:

Row vector

Column vector

$$\vec{V} = \begin{bmatrix} V_x & V_y & V_z \end{bmatrix} \quad \text{or} \quad \vec{V} = \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix} \tag{2.13}$$

The magnitude of a vector in three-dimensional Cartesian space is its length given by:

$$\left|\vec{V}\right| = \sqrt{V_x^2 + V_y^2 + V_z^2} \tag{2.14}$$

Vectors:

From the information in the previous slides we may then say that the unit vector in the direction of $\vec{V}$ is:

$$\hat{V} = \frac{\vec{V}}{\left|\vec{V}\right|} = \frac{V_x\hat{i} + V_y\hat{j} + V_z\hat{k}}{\sqrt{V_x^2 + V_y^2 + V_z^2}} = l\hat{i} + m\hat{j} + n\hat{k} \qquad (2.15)$$

Vectors:

In Physics the dimensions of a vector are usually limited to three. However in Mathematics and Computer Science this is not the case. A vector is then a list or set of numbers written in a row or column:

$$[V] = \begin{bmatrix} V_1 & V_2 & \dots & V_n \end{bmatrix} \quad \text{or} \quad [V] = \begin{bmatrix} V_1 \\ V_2 \\ \dots \\ V_n \end{bmatrix} \tag{2.16}$$

Where the respective representations of $\vec{V}$ above are as row and column vectors.

Addition and Subtraction of Vectors:

$$\vec{V} + \vec{U} = [V_i + U_i] = [V_1 + U_1, V_2 + U_2, \ldots, V_n + U_n \qquad (2.17)$$

$$\vec{V} - \vec{U} = [V_i - U_i] = [V_1 - U_1, V_2 - U_2, \ldots, V_n - U_n \qquad (2.18)$$

Multiplication by a Scalar:

$$\alpha \vec{V} = \alpha[V_i] = [\alpha V_1, \alpha V_2, \ldots, \alpha V_n] \qquad (2.19)$$

Transpose of a Vector:

If $\vec{V} = [V_1, V_2 \ \ldots \ V_n]$ then:

$$\vec{V}^T = \begin{bmatrix} V_1 \\ V_2 \\ \ldots \\ V_n \end{bmatrix} \qquad (2.20)$$

Multiplication of Two Vectors:

The ‘Dot’ Product or Scalar Product:

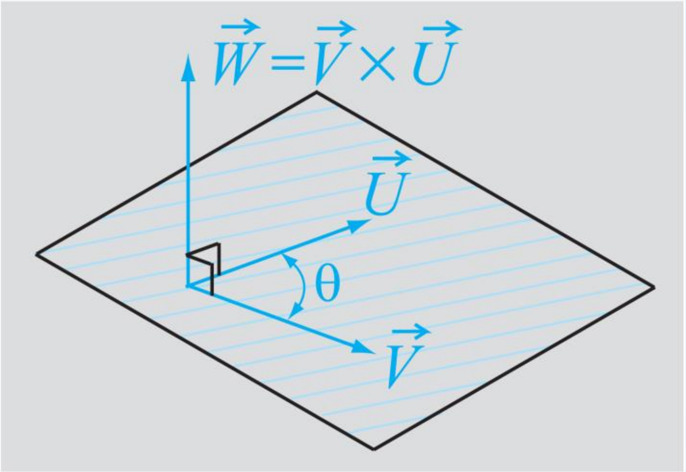
$$\vec{V} \bullet \vec{U} = [V_i][U_i] = V_1U_1 + V_2U_2 + \ldots V_nU_n \qquad (2.21)$$

or

$$\vec{V} \bullet \vec{U} = \left|\vec{V}\right|\left|\vec{U}\right| \cos\theta \qquad (2.23)$$

Multiplication of Two Vectors:

The ‘Cross’ Product or Vector Product:



$$\vec{W} = \vec{V} \times \vec{U} = \vec{V} \otimes \vec{U}$$

$$= |\vec{V}||\vec{U}| \sin\theta \, \hat{w}$$

$$= \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ V_x & V_y & V_z \\ U_x & U_y & U_z \end{vmatrix} = \begin{vmatrix} V_y & V_z \\ U_y & U_z \end{vmatrix} \hat{\imath} - \begin{vmatrix} V_x & V_z \\ U_x & U_z \end{vmatrix} \hat{\jmath} + \begin{vmatrix} V_x & V_y \\ U_x & U_y \end{vmatrix} \hat{k}$$

$$= (V_y U_z - V_z U_y)\hat{\imath} - (V_x U_z - V_z U_x)\hat{\jmath} + (V_x U_y - V_y U_x)\hat{k}$$

Linear Dependence and Independence of a Set of Vectors:

A set of vectors $\overrightarrow{V_1}, \overrightarrow{V_2}, \dots, \overrightarrow{V_n}$ is said to be linearly independent if

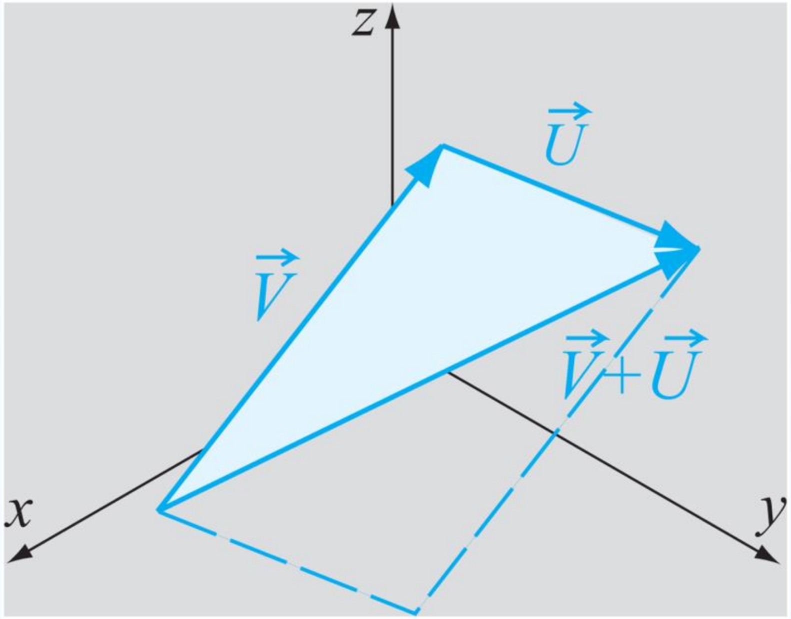
$$\alpha_1 \overrightarrow{V}_1 + \alpha_2 \overrightarrow{V}_2 + \dots + \alpha_n \overrightarrow{V}_n = 0 \qquad (2.26)$$

is satisfied $iff\ \alpha_1 = \alpha_2 = \dots = \alpha_n = 0$

Otherwise the set of vectors is said to be linearly dependent.

Another way of looking at linear independence/dependence is that if the vectors in the set cannot be written as a linear combination of any of the other vectors in the set then the set of vectors is linearly independent. The converse also holds true.

The Triangle Inequality:



$$\left|\vec{V}+\vec{U}\right| \leq \left|\vec{V}\right| + \left|\vec{U}\right| \tag{2.27}$$

Matrices and Linear Algebra

A matrix is a rectangular array of numbers. The size of a matrix refers to the number of rows and columns it contains. A matrix is denoted $[a]$ or A. The latter notation is used only where it is clear from the context that we are talking about a matrix.

$$[a] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{mn} \end{bmatrix} \tag{2.28}$$

An element of the matrix is referred to as a_{ij} where i and j refer to the row and column respectively where the element is located.
A row vector is simply a matrix with a single row or a $1 \times n$ matrix. A column vector is simply a matrix with a single column or an $n \times 1$ matrix

Multiplication by a Scalar:

If $[a] = [a_{ij}]$ is a matrix and α is a scalar then $\alpha[a] = [\alpha a_{ij}]$ is obtained by multiplying every element of $[a]$ by the number α.

Addition and Subtraction of Matrices:

Matrices can only be added or subtracted if they are both of the same size. This is done by adding or subtracting the corresponding entries of both matrices. The resulting matrix $[c]$ is then the same size where:

$$[c_{ij}] = [a_{ij}] + [b_{ij}] \tag{2.29}$$

and

$$[c_{ij}] = [a_{ij}] - [b_{ij}] \tag{2.30}$$

Transpose of a Matrix:

The transpose of a matrix is a matrix with the rows and columns interchanged. The first row becomes the first column of the transposed matrix and so on.

$$\begin{bmatrix} 2 & -1 & 0 \\ 5 & 3 & 1 \\ 6 & 1 & -4 \\ 7 & -2 & 9 \end{bmatrix}^T = \begin{bmatrix} 2 & 5 & 6 & 7 \\ -1 & 3 & 1 & -2 \\ 0 & 1 & -4 & 9 \end{bmatrix}$$

$$[a]^T = [a_{ij}^T] = [a_{ji}] \tag{2.31}$$

Multiplication of Matrices:

The multiplication $[c] = [a][b]$ is defined only if the number of columns of $[a]$ equals the number of rows of matrix $[b]$. The resulting matrix $[c]$ has the same number of rows as $[a]$ and the same number of columns as $[b]$.

$$[c]_{mn} = [a]_{mq}[b]_{qn}$$

same

So if $[a]$ is $m \times q$ and $[b]$ is $q \times n$ then $[c]$ is $m \times n$.

Multiplication of Matrices:

$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{bmatrix} \tag{2.32}$$

$$c_{11} = a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} + a_{14}b_{41} \tag{2.33}$$

$$c_{12} = a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} + a_{14}b_{42} \tag{2.34}$$

$$c_{21} = a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} + a_{24}b_{41} \tag{2.35}$$

etc.

In general:

$$c_{ij} = a_{ik}b_{kj} = \sum_{k=1}^{q} a_{ik}b_{kj} \tag{2.36}$$

Special Matrices:

Square Matrix

A matrix that has the same number of columns as rows is called a square matrix. Entries a_{ii} are known as the diagonal elements and all others the off-diagonal elements.

The elements above the diagonal $\left(a_{ij}; j > i\ \right)$ are called the above-diagonal entries or super-diagonal entries. The entries below the diagonal $\left(a_{ij}; i > j\right)$ are called the below-diagonal entries or sub-diagonal entries.

Example:

$$[A] = \begin{bmatrix} 4 & 5 & -3 \\ 8 & 1 & 2 \\ 4 & 6 & 3 \end{bmatrix}$$

is a square matrix

Special Matrices:

Diagonal Matrix

A square matrix with diagonal elements that are non-zero and off-diagonal elements that are all zero is called a diagonal matrix and is denoted $[D]$.

Example:

$$[D] = \begin{bmatrix} 6 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

is a 3×3 diagonal matrix.

Special Matrices:

Upper Triangular Matrix

A square matrix whose sub-diagonal entries are all zero is called an upper triangular matrix denoted $[U]$.

Example:

$$\begin{bmatrix} 1 & -3 & 1 \\ 0 & 5 & 12 \\ 0 & 0 & -7 \end{bmatrix}$$

is a 3×3 upper triangular matrix

Special Matrices:

Lower Triangular Matrix

A square matrix whose super-diagonal entries are all zero is called an lower triangular matrix denoted $[L]$.

Example:

$$\begin{bmatrix} 1 & 0 & 0 \\ 5 & 5 & 0 \\ 9 & -7 & -7 \end{bmatrix}$$

is a 3×3 lower triangular matrix

Special Matrices:

Identity Matrix

The identity matrix $[I]$ is a square matrix whose diagonal elements are all 1s and whose off-diagonal entries are all 0s. For matrices the identity matrix is the analog of the number 1. Any matrix that is multiplied by the identity matrix remains unchanged.

Example:

$$[I] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

In general

$$[a][I] = [a]$$

Special Matrices:

Zero Matrix

A zero matrix is a matrix whose entries are all zero. It is the analog of the number 0 in that any matrix multiplied by the zero matrix will yield a zero matrix.

Example:

$$[z] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$[z][a] = [z]$$

Special Matrices:

Symmetric Matrix

A symmetric matrix is a square matrix whose lower diagonal entries mirror its upper diagonal entries.

That is,

$$[a_{ij}] = [a_{ji}]$$

With symmetric matrices the transpose of the matrix is that matrix itself.

$$[a]^T = [a]$$

Example:

$$\begin{bmatrix} 3 & 4 & 8 \\ 4 & 1 & 2 \\ 8 & 2 & -7 \end{bmatrix}$$

is a symmetric matrix

Inverse of a Matrix:

A square matrix is invertible (i.e. its multiplicative inverse can be found) if there exists a square matrix $[b]$ of the same size such that $[a][b] = [I]$.

Put alternatively:

$$[a][b] = [I]$$

or

$$[a][a]^{-1} = [a]^{-1}[a] = [I]$$

where

$$[a]^{-1} = [b]$$

Example:

Example 2-1: Inverse of a matrix.

Show that the matrix $[b] = \begin{bmatrix} 0.1 & 0.2 & 0 \\ 0.4 & 0.1 & 0.2 \\ 0.2 & 0.1 & 0.8 \end{bmatrix}$ is the inverse of the matrix $[a] = \begin{bmatrix} -1.2 & 3.2 & -0.8 \\ 5.6 & -1.6 & 0.4 \\ -0.4 & -0.6 & 1.4 \end{bmatrix}$.

SOLUTION

To show that the matrix $[b]$ is the inverse of the matrix $[a]$, the two matrices are multiplied.

$$[a][b] = \begin{bmatrix} -1.2 & 3.2 & -0.8 \\ 5.6 & -1.6 & 0.4 \\ -0.4 & -0.6 & 1.4 \end{bmatrix} \begin{bmatrix} 0.1 & 0.2 & 0 \\ 0.4 & 0.1 & 0.2 \\ 0.2 & 0.1 & 0.8 \end{bmatrix} =$$

$$\begin{bmatrix} (1.2 \cdot 0.1 + 3.2 \cdot 0.4 + -0.8 \cdot 0.2) & (1.2 \cdot 0.2 + 3.2 \cdot 0.1 + -0.8 \cdot 0.1) & (1.2 \cdot 0 + 3.2 \cdot 0.2 + -0.8 \cdot 0.8) \\ (5.6 \cdot 0.1 + -1.6 \cdot 0.4 + 0.4 \cdot 0.2) & (5.6 \cdot 0.2 + -1.6 \cdot 0.1 + 0.4 \cdot 0.1) & (5.6 \cdot 0 + -1.6 \cdot 0.2 + 0.4 \cdot 0.8) \\ (-0.4 \cdot 0.1 + -0.6 \cdot 0.4 + 1.4 \cdot 0.2) & (-0.4 \cdot 0.2 + -0.6 \cdot 0.1 + 1.4 \cdot 0.1) & (-0.4 \cdot 0 + -0.6 \cdot 0.2 + 1.4 \cdot 0.8) \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Properties of Matrices:

- $[a] + [b] = [b] + [a]$
- $([a] \pm [b]) \pm [c] = [a] \pm ([b] \pm [c])$
- $\alpha([a] \pm [b]) = \alpha[a] \pm \alpha[b]$; where α is a scalar
- $(\alpha \pm \beta)[a] = \alpha[a] \pm \beta[a]$; where α and β are scalars
- In general $[a][b] \neq [b][a]$; unless one is the inverse of the other
- $([a] \pm [b])[c] = [a][c] \pm [b][c]$
- $[a]([b] \pm [c]) = [a][b] \pm [a][c]$
- $\alpha([a][b]) = (\alpha[a])[b] = [a](\alpha[b])$
- If $\exists\ [a][b]$ then $([a][b])^T = [b]^T[a]^T$;order of multiplication is changed
- $([a]^T)^T = [a]$
- $([a]^{-1})^{-1} = [a]$
- If $[a]$ and $[b]$ are two square invertible matrices of the same size then $([a][b])^{-1} = [b]^{-1}[a]^{-1}$; order of multiplication is changed

The Determinant of a Matrix:

Useful quantity that is used to determine if a matrix has an inverse.

$$\det(A) = |A| = \sum_{J} (-1)^k a_{1, j_1} a_{2, j_2} \ldots a_{n, j_n} \qquad (2.40)$$

- over all $n!$

Cramer's Rule and Solution to a System of Simultaneous Equations:

Given a set of n simultaneous equations with n unknowns:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \ldots \ + \ \ldots \ + \ldots + \ \ldots \ &= \ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n \end{aligned} \tag{2.41}$$

The system of equations can be written compactly using matrices (indeed this is the primary use of matrices!):

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \ldots \\ b_n \end{bmatrix} \tag{2.42}$$

i.e.

$$[a][x] = [b] \tag{2.43}$$

Cramer's Rule and Solution to a System of Simultaneous Equations:

Cramer's Rule states that the solution (for $[x]$), if it exists, is given by:

$$x_j = \frac{\det(a'_j)}{\det(a)} \text{ for } j = 1, 2, \ldots, n \tag{2.44}$$

where a'_j is the matrix formed by replacing the jth column of the matrix $[a]$ with the column vector $[b]$.

Note: $[a]^{-1}$ exists iff $\det(a) \neq 0$

$\det(a) = 0$ if one or more columns or rows of $[a]$ are not linearly independent.

Example:

Example 2-2: Solving a system of linear equations using Cramer's rule.

Find the solution of the following system of equations using Cramer's rule.

$$\begin{aligned} 2x + 3y - z &= 5 \\ 4x + 4y - 3z &= 3 \\ -2x + 3y - z &= 1 \end{aligned} \tag{2.45}$$

SOLUTION

Step 1: Write the system of equations in a matrix form $[a][x] = [b]$.

$$\begin{bmatrix} 2 & 3 & -1 \\ 4 & 4 & -3 \\ -2 & 3 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix} \tag{2.46}$$

Step 2: Calculate the determinant of the matrix of coefficients.

$$\begin{aligned} \det(A) &= 2[(4\times -1) - (-3\times 3)] - 3[(4\times -1) - (-3\times -2)] - 1[(4\times 3) - (4\times -2)] \\ &= 2(5) - 3(-10) - 1(20) = 10 + 30 - 20 = 20 \end{aligned}$$

Step 3: Apply Eq. (2.44) to find x, y, and z. To find x, the modified matrix a'_x is created by replacing its first column with $[b]$.

$$x = \frac{\det\left(\begin{bmatrix} 5 & 3 & -1 \\ 3 & 4 & -3 \\ 1 & 3 & -1 \end{bmatrix}\right)}{20} = \frac{(5\cdot 5) - (3\cdot 0) - (1\cdot 5)}{20} = 1$$

In the same way, to find y, the modified matrix a'_y is created by replacing its second column with $[b]$.

$$y = \frac{\det\left(\begin{bmatrix} 2 & 5 & -1 \\ 4 & 3 & -3 \\ -2 & 1 & -1 \end{bmatrix}\right)}{20} = \frac{(20\cdot 0) - (5\cdot -10) - (1\cdot 10)}{20} = 2$$

Finally, to determine the value of z, the modified matrix a'_z is created by replacing its third column with $[b]$.

$$z = \frac{\det\left(\begin{bmatrix} 2 & 3 & 5 \\ 4 & 4 & 3 \\ -2 & 3 & 1 \end{bmatrix}\right)}{20} = \frac{(2\cdot -5) - (3\cdot 10) - (5\cdot 20)}{20} = 3$$

To check the answer, the matrix of coefficients $[a]$ is multiplied by the solution:

$$\begin{bmatrix} 2 & 3 & -1 \\ 4 & 4 & -3 \\ -2 & 3 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2+6-3 \\ 4+8-9 \\ -2+6-3 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix}$$

The right-hand side is equal to $[b]$, which confirms that the solution is correct.

Ordinary Differential Equations (ODE):

An ordinary differential equation is one that contains one dependent variable (e.g. y), one independent variable (e.g. x) and ordinary (as opposed to partial) derivatives of the dependent variable.

It is said to be linear if its dependence on the dependent variable and its derivatives is linear.

Examples of linear ODEs:

$$\frac{dy}{dx} = 10x$$

$$c\frac{dx}{dt} + kx = -m\frac{d^2x}{dt^2}$$

Ordinary Differential Equations (ODE):

An ODE is non-linear if any of the coefficients are functions of the dependent variable.

Examples:

$$\frac{d^2y}{dt^2} + \sin y = 4$$

$$y\frac{d^2y}{dt^2} + 3y = 8$$

Ordinary Differential Equations (ODE):

An ODE is said to be homogeneous if the coefficient of the independent variables is zero. Otherwise it is said to be nonhomogeneous.

Examples:

$$\frac{dy}{dx} = 10x$$

Is a linear, nonhomogeneous ODE.

$$\frac{d^2y}{dt^2} + \sin y = 4$$

Is a non-linear, homogeneous ODE.

Ordinary Differential Equations (ODE):

The order of an ODE is the highest derivative that appears in the equation.

Example:

$$\frac{d^2y}{dt^2} + \sin y = 4$$

Is a second order ODE.

Ordinary Differential Equations (ODE):

Boundary Conditions/Initial Conditions:

To eliminate integration constants when solving a DE we apply constraints. These are referred to as boundary conditions (solutions to the DE at particular points) and in the case of time dependent DEs, initial conditions (or the solution to the DE at $t = 0$).

Solution to Non-Homogeneous Linear First Order ODE:

Consider:

$$\frac{dy}{dx} + P(x)y = Q(x) \qquad (2.48)$$

Multiply both sides of this equation by the following integrating factor:

$$\mu(x) = e^{\int P(x)dx} \qquad (2.49)$$

This gives (do it as an exercise!):

$$\frac{d}{dx}(y\mu) = Q(x)\mu(x) \qquad (2.50)$$

Solution to a Non-Homogeneous Linear First-Order ODE:

Integrating both sides gives:

$$y(x)\mu(x) = \int Q(x)\mu(x)dx + C_1 \qquad (2.51)$$

Dividing through by $\mu(x)$ gives:

$$y(x) = \frac{1}{\mu(x)}\int Q(x)\mu(x)dx + \frac{C_1}{\mu(x)} \qquad (2.52)$$

The constant of integration is determined from a constraint which is problem dependent.

Solution to an Homogeneous Linear Second-Order ODE:

Consider:

$$\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0 \qquad (2.53)$$

where a and b are constants.

The general solution to this equation is found by substituting $y = e^{sx}$. The resulting equation is called the characteristic equation:

$$s^2 + bs + c = 0 \qquad (2.54)$$

Solution to an Homogeneous Linear Second-Order ODE:

The solution is obtained from the quadratic formula:

$$s = \frac{-b \pm \sqrt{b^2 - 4c}}{2} \qquad (2.55)$$

The general solution is then:

$$y(x) = e^{-bx/2}\left[C_1 e^{\frac{x}{2}\sqrt{b^2-4c}} + C_2 e^{-\frac{x}{2}\sqrt{b^2-4c}}\right] \qquad (2.56)$$

If $b^2 < 4ac$ we get:

$$y(x) = e^{-bx/2}\left[C_1 e^{\frac{i}{2}x\sqrt{b^2-4c}} + C_2 e^{-\frac{i}{2}x\sqrt{b^2-4c}}\right] \qquad (2.57)$$

Functions of Two or More Independent Variables:

Consider the function:

$$z = f(x, y) = \frac{x^2}{4} + \frac{y^2}{9}$$

This is a function of two variables where z is the dependent variable and x and y are the independent variables.

The Partial Derivative:

For $z = f(x, y)$, the first partial derivative of f with respect to x is denoted by $\frac{\delta f}{\delta x}$ or f_x and is defined by:

$$\frac{\partial f}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x} \qquad (2.61)$$

if the limit exists. Also:

$$\frac{\partial f}{\partial y} = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y} \qquad (2.62)$$

The Partial Derivative:

The partial derivative with respect to a certain variable (say x) simply by treating (or 'holding') all other independent variables (say y and z) as constants.

The ***total differential*** (or exact differential) of a function of two variables, say $f(x, y)$, is given by:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy \qquad (2.63)$$

The Partial Derivative:

Using the formula for the total differential we can state:

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \qquad (2.64)$$

where x and y are both functions of the independent variable t

The Partial Derivative:

Consider $f(x, y)$ where $y = g(x)$ then:

$$\frac{df}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial x} \qquad (2.67)$$

Consider $f(x, y, z)$ where $z = h(x, y)$ then:

$$\left.\frac{\partial f}{dx}\right|_y = \frac{\partial f}{\partial x} + \left.\frac{\partial f}{\partial z}\frac{\partial z}{\partial x}\right|_y \qquad (2.68)$$

or

$$\frac{\partial f}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x} \qquad (2.69)$$

The Jacobian:

The Jacobian is a quantity that arises when solving systems of non-linear equations. Given $f_1(x, y) = a$ and $f_2(x, y) = b$ (where a and b are constants) then the Jacobian matrix is:

$$[J] = \begin{bmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{bmatrix} \qquad (2.71)$$

The Jacobian determinant or simply the Jacobian is:

$$f_1, f_2) = \det\left(\begin{vmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{vmatrix}\right) = \left(\frac{\partial f_1}{\partial x}\right)\left(\frac{\partial f_2}{\partial y}\right) - \left(\frac{\partial f_1}{\partial y}\right)\left(\frac{\partial}{\partial}\right. \qquad (2.72)$$

The Jacobian:

In general the Jacobian is:

$$J(f_1, f_2, \ldots, f_n) = \det\left(\begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \cdots & \dfrac{\partial f_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial f_n}{\partial x_1} & \dfrac{\partial f_n}{\partial x_2} & \cdots & \dfrac{\partial f_n}{\partial x_n} \end{bmatrix}\right) \tag{2.73}$$

More about Jacobians later in the course!

Taylor Series Expansion of Functions:

It is used to represent a function as an infinite power series. This can be useful when trying to integrate certain non-analytic functions since a truncated series may result in a good analytic approximation to the function.

It may also be used in approximating functions to make them more easily manipulated for certain purposes such as increasing computational speed.

Given a function $f(x)$ that is differentiable $n+1$ times in an interval containing $x = x_0$, Taylor's theorem states that

$\exists\ \xi\ \in [x, x_0]$ such that:

$$\begin{aligned} f(x) = f(x_0) + (x - x_0)\left.\frac{df}{dx}\right|_{x = x_0} + \frac{(x-x_0)^2}{2!}\left.\frac{d^2 f}{dx^2}\right|_{x = x_0} + \frac{(x-x_0)^3}{3!}\left.\frac{d^3 f}{dx^3}\right|_{x = } \\ + \ldots + \frac{(x-x_0)^n}{n!}\left.\frac{d^n f}{dx^n}\right|_{x = x_0} + R_n(x) \end{aligned} \tag{2.74}$$

Taylor Series Expansion of Functions:

Where

$$R_n = \frac{(x-x_0)^{n+1}}{(n+1)!} \frac{d^{n+1} f}{dx^{n+1}}\bigg|_{x=\xi} \qquad (2.75)$$

Note that R_n cannot be calculated since ξ is unknown.

For $n = 0$ the Taylor Series reduces to:

$$f(x) = f(x_0) + (x-x_0)\frac{df}{dx}\bigg|_{x=\xi} \quad \text{or} \quad \frac{df}{dx}\bigg|_{x=\xi} = \frac{f(x)-f(x_0)}{(x-x_0)} \qquad (2.76)$$

which is a statement of the Mean Value Theorem.

Example:

Example 2-3: Approximation of a function with Taylor series expansion.

Approximate the function $y = \sin(x)$ by using Taylor series expansion about $x = 0$, using two, four, and six terms.

(*a*) In each case, calculate the approximate value of the function at $x = \frac{\pi}{12}$, and at $x = \frac{\pi}{2}$.

(*b*) Using MATLAB, plot the function and the three approximations for $0 \le x \le \pi$.

SOLUTION

The first five derivatives of the function $y = \sin(x)$ are:

$$y' = \cos(x), \quad y'' = -\sin(x), \quad y^{(3)} = -\cos(x), \quad y^{(4)} = \sin(x), \quad \text{and} \quad y^{(5)} = \cos(x)$$

At $x = 0$, the values of these derivatives are:

$$y' = 1, \quad y'' = 0, \quad y^{(3)} = -1, \quad y^{(4)} = 0, \quad \text{and} \quad y^{(5)} = 1$$

Substituting this information and $y(0) = \sin(0) = 0$ in Eq. (2.74) gives:

$$y(x) = 0 + x + 0 - \frac{x^3}{3!} + 0 + \frac{x^5}{5!} \tag{2.77}$$

(*a*) For $x = \frac{\pi}{12}$, the exact value of the function is $y = \sin\left(\frac{\pi}{12}\right) = \frac{1}{4}(\sqrt{6} + \sqrt{2}) = 0.2588190451$

The approximate values using two, four, and six terms of the Taylor series expansion are:

Using two terms in Eq. (2.77) gives: $y(x) = x = \frac{\pi}{12} = 0.2617993878$

Using four terms in Eq. (2.77) gives: $y(x) = x - \frac{x^3}{3!} = \frac{\pi}{12} - \frac{(\pi/12)^3}{3!} = 0.2588088133$

Using six terms in Eq. (2.77) gives: $y(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} = \frac{\pi}{12} - \frac{(\pi/12)^3}{3!} + \frac{(\pi/12)^5}{5!} = 0.2588190618$

For $x = \frac{\pi}{2}$, the exact value of the function is $y = \sin\left(\frac{\pi}{2}\right) = 1$

Using two terms in Eq. (2.77) gives: $y(x) = x = \frac{\pi}{2} = 1.570796327$

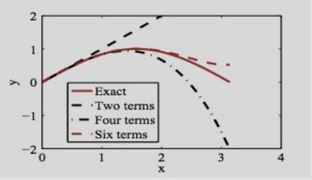
Using four terms in Eq. (2.77) gives: $y(x) = x - \frac{x^3}{3!} = \frac{\pi}{2} - \frac{(\pi/2)^3}{3!} = 0.9248322293$

Using six terms in Eq. (2.77) gives: $y(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} = \frac{\pi}{2} - \frac{(\pi/2)^3}{3!} + \frac{(\pi/2)^5}{5!} = 1.004524856$

(*b*) Using a MATLAB program, listed in the following script file, the function and the three approximations were calculated for the domain $0 \le x \le \pi$. The program also plots the results.

```
x = linspace(0,pi,40);
y = sin(x);
y2 = x;
y4 = x - x.^3/factorial(3);
y6 = x - x.^3/factorial(3) + x.^5/factorial(5);
plot(x,y,'r',x,y2,'k--',x,y4,'k-.',x,y6,'r--')
axis([0,4,-2,2])
legend('Exact','Two terms','Four terms','Six terms')
xlabel('x'); ylabel('y')
```

The plot produced by the program is shown on the right. The results from both parts show, as expected, that the approximation of the function with the Taylor series is more accurate when more terms are used and when the point at which the value of the function is desired is close to the point about which the function is expanded.



Taylor Series for a Function of Two Variables:

Taylor's expansion for a function of two variables is done the same way as for a function of one independent variable except that the differentiation involves partial derivatives.

$$
\begin{aligned}
f(x, y) &= f(x_0, y_0) + \frac{1}{1!}\left[(x - x_0)\frac{\partial f}{\partial x}\bigg|_{x_0, y_0} + (y - y_0)\frac{\partial f}{\partial y}\bigg|_{x_0, y_0}\right] + \\
&\frac{1}{2!}\left[(x - x_0)^2\frac{\partial^2 f}{\partial x^2}\bigg|_{x_0, y_0} + 2(x - x_0)(y - y_0)\frac{\partial^2 f}{\partial x \partial y}\bigg|_{x_0, y_0} + (y - y_0)^2\frac{\partial^2 f}{\partial y^2}\bigg|_{x_0, y_0}\right] + \\
&+ \ldots + \frac{1}{n!}\left[\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}(x - x_0)^k(y - y_0)^{n-k}\frac{\partial^n f}{\partial x^k \partial y^{n-k}}\bigg|_{x_0, y_0}\right]
\end{aligned}
\tag{2.78}
$$

Inner Product and Orthogonality:

The inner product (scalar product or 'dot' product)of two vectors $\vec{V}$ and $\vec{U}$ is denoted $\langle\vec{V}|\vec{U}\rangle$ or $\langle\vec{V},\vec{U}\rangle$ or $\vec{V}\cdot\vec{U}$.

Say $\vec{V} = V_x\hat{\imath} + V_y\hat{\jmath} + V_z\hat{k}$ and $\vec{U} = U_x\hat{\imath} + U_y\hat{\jmath} + U_z\hat{k}$.

Then:

$$\langle \vec{V} \mid \vec{U} \rangle = \vec{V} \bullet \vec{U} = \left|\vec{V}\right|\left|\vec{U}\right| \cos\theta = V_x U_x + V_y U_y + V_z U_z \qquad (2.80)$$

The vectors are said to be orthogonal to each other (or simply orthogonal) if:

$$\langle\vec{V}|\vec{U}\rangle = 0$$

and parallel if: $\langle\vec{V}|\vec{U}\rangle = 1.$

Inner Product and Orthogonality:

The inner product of two functions $f(x)$ and $g(x)$ over an interval $[a, b]$ is given by:

$$\langle f(x)|g(x)\rangle = \int_a^b f(x)g(x)dx$$

Again, they are said to be orthogonal if:

$$\langle f(x)|g(x)\rangle = 0$$

Importantly the sine and cosine functions are orthogonal at all frequencies. That is:

$$\langle \sin(kx)|\cos(mx)\rangle = \int_{-\pi}^{\pi} \sin(kx)\cos(mx)dx = 0 \quad \text{for both } k = m \text{ and } k \neq m \tag{2.83}$$

Recommended Problems:

Problems to be solved by hand (do at least 2):
2.2, 2.8, 2.22

Problems to be programmed in Matlab (do at least 1):
2.27, 2.31

Problems in math, science and engineering (do at least 1):

2.32, 2.34, 2.36(a)

You should pick out problems that you find interesting/challenging and do these too. Use SI units (without conversion).