VECTOR CALCULUS

Vector calculus is a field of mathematics concerned with multivariate real analysis of vectors in two or more dimensions. It consists of set of problems solving techniques very useful for engineering and physics.

SCALAR AND VECTOR POINT FUNCTIONS

Let $\vec{F}(t) = f_1(t)\hat{\imath} + f_2(t)\hat{\jmath} + f_3(t)\hat{k}$ be a 'Vector function'. Then for various values of t we get a set of constant vectors.

Let $\varphi = \varphi(x, y, z)$ be a 'Scalar function'. Then for various values of x, y, z we get a set of points or scalars.

Vector operation $\nabla(del)$ is defined by the equation

$$\nabla = \frac{\partial}{\partial x}\hat{\imath} + \frac{\partial}{\partial y}\hat{\jmath} + \frac{\partial}{\partial z}\hat{k}$$

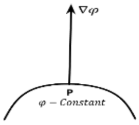
This operator has a great role in vector calculus. Laplacian operator ∇^2 is defined as follows

$$\nabla^2 = \nabla.\nabla = \left(\frac{\partial}{\partial x}\hat{\imath} + \frac{\partial}{\partial y}\hat{\jmath} + \frac{\partial}{\partial z}\hat{k}\right).\left(\frac{\partial}{\partial x}\hat{\imath} + \frac{\partial}{\partial y}\hat{\jmath} + \frac{\partial}{\partial z}\hat{k}\right)$$

$$\boldsymbol{\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}}$$

<u>Gradient</u>

The vector function $\nabla\varphi$ is defined as the gradient of the scalar function $\varphi = \varphi(x, y, z)$



$i.e., grad\varphi = \nabla\varphi = \left(\frac{\partial}{\partial x}\hat{\imath} + \frac{\partial}{\partial y}\hat{\jmath} + \frac{\partial}{\partial z}\hat{k}\right)\varphi$

$$\boldsymbol{grad\varphi = \nabla\varphi = \left(\frac{\partial\varphi}{\partial x}\hat{\imath} + \frac{\partial\varphi}{\partial y}\hat{\jmath} + \frac{\partial\varphi}{\partial z}\hat{k}\right)}$$

Geometrically, $\nabla\varphi$ represents a normal at any point P to the surface $\varphi(x, y, z) = constant$ and has a magnitude equal to the rate of change of $\varphi(x, y, z)$ along this normal. $\nabla\varphi$ is a **vector quantity**.

Note:

1. The unit normal vector $\hat{n}$ along $\nabla\varphi$ is given by $\boldsymbol{\hat{n} = \frac{\nabla\varphi}{|\nabla\varphi|}}$ or $\hat{n} = \frac{\nabla f}{|\nabla f|}$

2. The component of $\nabla\varphi$ in the direction of a unit vector $\vec{a}$ is $\nabla\varphi.\hat{n}$ and is called the *directional derivative* of φ in the direction of $\vec{a}$. Thus, the directional derivative is maximum in the direction $\nabla\varphi$ and the magnitude of this maximum is equal to $|\nabla\varphi|$.

$i.e.,$ $\boldsymbol{D.D = \nabla\varphi.\hat{n}}$ where $\hat{n} = \frac{\vec{a}}{|\vec{a}|}$

Problems

1. Find the unit normal vector to the surface $x^3+y^3+3xyz=3$ at $(1,2,-1)$.

Sol: Let $\varphi=x^3+y^3+3xyz-3$

$\frac{\partial\varphi}{\partial x}=3x^2+3yz$ $\quad\frac{\partial\varphi}{\partial y}=3y^2+3xz$ $\quad\frac{\partial\varphi}{\partial z}=3xy$

Now, $\nabla\varphi=\left(\frac{\partial\varphi}{\partial x}\hat{\imath}+\frac{\partial\varphi}{\partial y}\hat{\jmath}+\frac{\partial\varphi}{\partial z}\hat{k}\right)$

$$\nabla\varphi=(3x^2+3yz)\hat{\imath}+(3y^2+3xz)\hat{\jmath}+(3xy)\hat{k}$$

At $(1,2,-1)$

$$\nabla\varphi=(3-6)\hat{\imath}+(12-3)\hat{\jmath}+(6)\hat{k}$$

$$\nabla\varphi=-3\hat{\imath}+9\hat{\jmath}+6\hat{k}$$

$$|\nabla\varphi|=\sqrt{(-3)^2+(9)^2+(6)^2}=\sqrt{9+81+36}$$

$$|\boldsymbol{\nabla\varphi}|=\sqrt{\mathbf{126}}$$

The unit normal vector , $\hat{\boldsymbol{n}}=\frac{\boldsymbol{\nabla\varphi}}{|\boldsymbol{\nabla\varphi}|}=\frac{-3\hat{\imath}+9\hat{\jmath}+6\hat{k}}{\sqrt{\mathbf{126}}}$

2. Find the unit normal vector to the surface $x^2y+y^2z+z^2x=5$ at $(1,-1,2)$.

Sol: Let $\varphi=x^2y+y^2z+z^2x-5$

$\frac{\partial\varphi}{\partial x}=2xy+z^2.1$ $\quad\frac{\partial\varphi}{\partial y}=x^2.1+2yz$ $\quad\frac{\partial\varphi}{\partial z}=y^2+2zx$

Now, $\nabla\varphi=\left(\frac{\partial\varphi}{\partial x}\hat{\imath}+\frac{\partial\varphi}{\partial y}\hat{\jmath}+\frac{\partial\varphi}{\partial z}\hat{k}\right)$

$$\nabla\varphi=(2xy+z^2)\hat{\imath}+(x^2+2yz)\hat{\jmath}+(y^2+2zx)\hat{k}$$

At $(1,-1,2)$

$$\nabla\varphi=(-2+4)\hat{\imath}+(1-4)\hat{\jmath}+(1+4)\hat{k}$$

$$\boldsymbol{\nabla\varphi}=\mathbf{2}\hat{\boldsymbol{\imath}}-\mathbf{3}\hat{\boldsymbol{\jmath}}+\mathbf{5}\hat{\boldsymbol{k}}$$

$$|\nabla\varphi|=\sqrt{(2)^2+(-3)^2+(5)^2}=\sqrt{4+9+25}$$

$$|\boldsymbol{\nabla\varphi}|=\sqrt{\mathbf{38}}$$

The unit normal vector , $\hat{\boldsymbol{n}}=\frac{\boldsymbol{\nabla\varphi}}{|\boldsymbol{\nabla\varphi}|}=\frac{2\hat{\imath}-3\hat{\jmath}+5\hat{k}}{\sqrt{\mathbf{38}}}$

3. Find the directional derivative of the function $\varphi = xy^2 + yz^3$ at $(2,-1,1)$ along $\hat{\imath} + 2\hat{\jmath} + 3\hat{k}$

Sol: Given $\varphi = xy^2 + yz^3$ Let $\vec{a} = \hat{\imath} + 2\hat{\jmath} + 3\hat{k}$

$\frac{\partial\varphi}{\partial x} = 1.y^2$ $\frac{\partial\varphi}{\partial y} = x.2y + 1.z^3$ $\frac{\partial\varphi}{\partial z} = 0 + y.3z^2$

$\frac{\partial\varphi}{\partial x} = y^2$ $\frac{\partial\varphi}{\partial y} = 2xy + z^3$ $\frac{\partial\varphi}{\partial z} = 3yz^2$

Now, $\nabla\varphi = \left(\frac{\partial\varphi}{\partial x}\hat{\imath} + \frac{\partial\varphi}{\partial y}\hat{\jmath} + \frac{\partial\varphi}{\partial z}\hat{k}\right)$

$$\nabla\varphi = (y^2)\hat{\imath} + (2xy + z^3)\hat{\jmath} + (3yz^2)\hat{k}$$

At $(2,-1,1)$

$$\nabla\varphi = (1)\hat{\imath} + (-4+1)\hat{\jmath} + (-3)\hat{k}$$

$$\boldsymbol{\nabla\varphi = \hat{\imath} - 3\hat{\jmath} - 3\hat{k}}$$

Also, $\boldsymbol{\vec{a} = \hat{\imath} + 2\hat{\jmath} + 3\hat{k}}$

$$|\vec{a}| = \sqrt{1^2 + 2^2 + 3^2}$$

$$|\vec{a}| = \sqrt{1+4+9}$$

$$\boldsymbol{|\vec{a}| = \sqrt{14}}$$

$\therefore \hat{n} = \frac{\vec{a}}{|\vec{a}|} = \frac{\boldsymbol{\hat{\imath}+2\hat{\jmath}+3\hat{k}}}{\boldsymbol{\sqrt{14}}}$

$$\boldsymbol{D.D = \nabla\varphi.\hat{n}}$$

$$D.D = (\hat{\imath} - 3\hat{\jmath} - 3\hat{k}).\frac{(\hat{\imath} + 2\hat{\jmath} + 3\hat{k})}{\sqrt{14}}$$

$$D.D = \frac{(1)(1) + (-3)(2) + (-3)(3)}{\sqrt{14}}$$

$$D.D = \frac{1-6-9}{\sqrt{14}}$$

$$D.D = \frac{-14}{\sqrt{14}}$$

$$\boldsymbol{D.D = -\sqrt{14}}$$

4. Find the directional derivative of the function $f = 4xz^3 - 3x^2y^2z$ at $(2,-1,2)$ along $2\hat{\imath} - 3\hat{\jmath} + 6\hat{k}$

Sol: Given $f = 4xz^3 - 3x^2y^2z$ Let $\vec{a} = 2\hat{\imath} - 3\hat{\jmath} + 6\hat{k}$

$\frac{\partial f}{\partial x} = 4z^3.1 - 3y^2z.2x$ $\frac{\partial f}{\partial y} = 0 - 3x^2z.2y$ $\frac{\partial f}{\partial z} = 4x.3z^2 - 3x^2y^2.1$

$\frac{\partial f}{\partial x} = 4z^3 - 6xy^2z$ $\frac{\partial f}{\partial y} = -6x^2yz$ $\frac{\partial f}{\partial z} = 12xz^2 - 3x^2y^2$

Now, $\nabla f = \left(\frac{\partial f}{\partial x}\hat{\imath} + \frac{\partial f}{\partial y}\hat{\jmath} + \frac{\partial f}{\partial z}\hat{k}\right)$

$$\nabla f = (4z^3 - 6xy^2z)\hat{\imath} + (-6x^2yz)\hat{\jmath} + (12xz^2 - 3x^2y^2)\hat{k}$$

At $(2,-1,2)$

$$\nabla f = (32 - 24)\hat{\imath} + (48)\hat{\jmath} + (96 - 12)\hat{k}$$

$$\boldsymbol{\nabla f = 8\hat{\imath} + 48\hat{\jmath} + 84\hat{k}}$$

Also, $\boldsymbol{\vec{a} = 2\hat{\imath} - 3\hat{\jmath} + 6\hat{k}}$

$$|\vec{a}| = \sqrt{2^2 + (-3)^2 + 6^2}$$

$$|\vec{a}| = \sqrt{4 + 9 + 36}$$

$$|\vec{a}| = \sqrt{49}$$

$$\boldsymbol{|\vec{a}| = 7}$$

$\therefore \hat{n} = \frac{\vec{a}}{|\vec{a}|} = \frac{\boldsymbol{2\hat{\imath} - 3\hat{\jmath} + 6\hat{k}}}{\boldsymbol{7}}$

$$\boldsymbol{D.D = \nabla f.\hat{n}}$$

$$D.D = (8\hat{\imath} + 48\hat{\jmath} + 84\hat{k}).\frac{(2\hat{\imath} - 3\hat{\jmath} + 6\hat{k})}{7}$$

$$D.D = \frac{(8)(2) + (48)(-3) + (84)(6)}{7}$$

$$D.D = \frac{16 - 144 + 504}{7}$$

$$\boldsymbol{D.D = \frac{376}{7}}$$

5. Find the directional derivative of $\varphi = e^{2x}\cos(yz)$ at the origin in the direction of the tangent to the curve $x = asint, y = acost$ and $z = at$ at $t = \frac{\pi}{4}$.

Sol: Given $\varphi = e^{2x}\cos(yz)$

$\frac{\partial\varphi}{\partial x} = \cos(yz)\,.\,2e^{2x}$ $\qquad$ $\frac{\partial\varphi}{\partial y} = e^{2x}[-\sin(yz)\,.\,z]$ $\qquad$ $\frac{\partial\varphi}{\partial z} = e^{2x}[-\sin(yz)\,.\,y]$

$\frac{\partial\varphi}{\partial x} = 2e^{2x}\cos(yz)$ $\qquad$ $\frac{\partial\varphi}{\partial y} = -e^{2x}z\sin(yz)$ $\qquad$ $\frac{\partial\varphi}{\partial z} = -e^{2x}y\sin(yz)$

Now, $\nabla\varphi = \left(\frac{\partial\varphi}{\partial x}\hat{\imath} + \frac{\partial\varphi}{\partial y}\hat{\jmath} + \frac{\partial\varphi}{\partial z}\hat{k}\right)$

$$\nabla\varphi = [2e^{2x}\cos(yz)]\hat{\imath} + [-e^{2x}z\sin(yz)]\hat{\jmath} + [-e^{2x}y\sin(yz)\,]\hat{k}$$

At origin i.e., (0,0,0)

$\nabla\varphi = (2)\hat{\imath} + (0)\hat{\jmath} + (0)\hat{k}$ $\qquad$ $(since\ e^0 = \cos(0) = 1\,, \sin(0) = 0)$

$\nabla\varphi = 2\hat{\imath}$

Consider, $\vec{r} = x\hat{\imath} + y\hat{\jmath} + z\hat{k}$

$$\vec{r} = (a\sin t)\hat{\imath} + (a\cos t)\hat{\jmath} + (at)\,\hat{k}$$

$$\frac{d\vec{r}}{dt} = (a\cos t)\hat{\imath} + (-a\sin t)\hat{\jmath} + (a)\,\hat{k}$$

At $t = \frac{\pi}{4}$

$$\frac{d\vec{r}}{dt} = \left(a.\frac{1}{\sqrt{2}}\right)\hat{\imath} + \left(-a.\frac{1}{\sqrt{2}}\right)\hat{\jmath} + (a)\hat{k}$$

$$\frac{d\vec{r}}{dt} = \left(\frac{a}{\sqrt{2}}\right)\hat{\imath} - \left(\frac{a}{\sqrt{2}}\right)\hat{\jmath} + (a)\hat{k}$$

Unit normal vector , $\hat{n} = \frac{\frac{d\vec{r}}{dt}}{\left|\frac{d\vec{r}}{dt}\right|}$

$$\hat{n} = \frac{\left(\frac{a}{\sqrt{2}}\right)\hat{\imath} - \left(\frac{a}{\sqrt{2}}\right)\hat{\jmath} + (a)\hat{k}}{\sqrt{\left(\frac{a}{\sqrt{2}}\right)^2 + \left(-\frac{a}{\sqrt{2}}\right)^2 + a^2}}$$

$$\hat{n} = \frac{\left(\frac{a}{\sqrt{2}}\right)\hat{\imath} - \left(\frac{a}{\sqrt{2}}\right)\hat{\jmath} + (a)\hat{k}}{\sqrt{\frac{a^2}{2} + \frac{a^2}{2} + a^2}} = \frac{\left(\frac{a}{\sqrt{2}}\right)\hat{\imath} - \left(\frac{a}{\sqrt{2}}\right)\hat{\jmath} + (a)\hat{k}}{\sqrt{a^2 + a^2}}$$

$$\hat{n} = \frac{\left(\frac{a}{\sqrt{2}}\right)\hat{\imath} - \left(\frac{a}{\sqrt{2}}\right)\hat{\jmath} + (a)\hat{k}}{\sqrt{2}.\,a}$$

$$\hat{n} = \frac{\frac{a}{\sqrt{2}}\left(\hat{\imath} - \hat{\jmath} + \sqrt{2}\hat{k}\right)}{\sqrt{2}.a}$$

$$\hat{n} = \frac{\left(\hat{\imath} - \hat{\jmath} + \sqrt{2}\hat{k}\right)}{2}$$

$$\boldsymbol{D.D = \nabla\varphi.\hat{n}}$$

$D.D = 2\hat{\imath}.\frac{(\hat{\imath}-\hat{\jmath}+\sqrt{2}\hat{k})}{2}$

$\boldsymbol{D.D} = \mathbf{1}$

6. Find the directional derivative of $\varphi = xy^2 + yz^3$ at the point $(2,-1,1)$ in the direction normal to the surface $xlogz - y^2 + 4$ at the point $(-1,2,1)$.

Sol: Given $\varphi = xy^2 + yz^3$

$\frac{\partial\varphi}{\partial x} = 1.y^2$ $\quad$ $\frac{\partial\varphi}{\partial y} = x.2y + 1.z^3$ $\quad$ $\frac{\partial\varphi}{\partial z} = 0 + y.3z^2$

$\frac{\partial\varphi}{\partial x} = y^2$ $\quad$ $\frac{\partial\varphi}{\partial y} = 2xy + z^3$ $\quad$ $\frac{\partial\varphi}{\partial z} = 3yz^2$

Now, $\nabla\varphi = \left(\frac{\partial\varphi}{\partial x}\hat{\imath} + \frac{\partial\varphi}{\partial y}\hat{\jmath} + \frac{\partial\varphi}{\partial z}\hat{k}\right)$

$$\nabla\varphi = (y^2)\hat{\imath} + (2xy + z^3)\hat{\jmath} + (3yz^2)\hat{k}$$

At $(2,-1,1)$

$$\nabla\varphi = (1)\hat{\imath} + (-4 + 1)\hat{\jmath} + (-3)\hat{k}$$

$$\nabla\varphi = \hat{\imath} - 3\hat{\jmath} - 3\hat{k}$$

Also given $\psi = xlogz - y^2 + 4$

$\frac{\partial\psi}{\partial x} = 1.logz - 0$ $\quad$ $\frac{\partial\psi}{\partial y} = 0 - 2y$ $\quad$ $\frac{\partial\psi}{\partial z} = x.\frac{1}{z} - 0$

$\frac{\partial\psi}{\partial x} = logz$ $\quad$ $\frac{\partial\psi}{\partial y} = -2y$ $\quad$ $\frac{\partial\psi}{\partial z} = \frac{x}{z}$

Now, $\nabla\psi = \left(\frac{\partial\psi}{\partial x}\hat{\imath} + \frac{\partial\psi}{\partial y}\hat{\jmath} + \frac{\partial\psi}{\partial z}\hat{k}\right)$

$$\nabla\psi = (logz)\hat{\imath} + (-2y)\hat{\jmath} + \left(\frac{x}{z}\right)\hat{k}$$

At $(-1,2,1)$

$$\nabla\psi = (log1)\hat{\imath} + (-2.2)\hat{\jmath} + \left(\frac{-1}{1}\right)\hat{k}$$

$$\nabla\psi = (0)\hat{\imath} + (-4)\hat{\jmath} + (-1)\hat{k}$$

$$|\nabla\psi| = \sqrt{(-4)^2 + (-1)^2}$$

$$\boldsymbol{|\nabla\psi| = \sqrt{17}}$$

$$\hat{\boldsymbol{n}} = \frac{\nabla\psi}{|\nabla\psi|} = \frac{0\hat{\imath} - 4\hat{\jmath} - \hat{k}}{\sqrt{17}}$$

$$\boldsymbol{D.D = \nabla\varphi.\hat{n}}$$

$$D.D = \nabla\varphi.\frac{\nabla\psi}{|\nabla\psi|}$$

$$D.D = (\hat{\imath}\ - 3\hat{\jmath} - 3\hat{k}).\frac{(0\hat{\imath} - 4\hat{\jmath} - \hat{k})}{\sqrt{17}}$$

$$D.D = \frac{(1)(0) + (-3)(-4) + (-3)(-1)}{\sqrt{17}}$$

$$D.D = \frac{0 + 12 + 3}{\sqrt{17}}$$

$$\boldsymbol{D.D = \frac{15}{\sqrt{17}}}$$

7. The directional derivative of $\varphi = axy^2 + byz + cz^2x^3$ at the point $(-1,1,2)$ has maximum magnitude of 32 units in the direction parallel to $y - axis$ find a, b, c.

Sol: Given $\varphi = axy^2 + byz + cz^2x^3$

$\frac{\partial\varphi}{\partial x} = ay^2.1 + 0 + cz^2.3x^2$ $\quad$ $\frac{\partial\varphi}{\partial y} = ax.2y + bz.1 + 0$ $\quad$ $\frac{\partial\varphi}{\partial z} = 0 + by.1 + cx^3.2z$

$\frac{\partial\varphi}{\partial x} = ay^2 + 3cz^2x^2$ $\quad$ $\frac{\partial\varphi}{\partial y} = 2axy + bz$ $\quad$ $\frac{\partial\varphi}{\partial z} = by + 2czx^3$

Now, $\nabla\varphi = \left(\frac{\partial\varphi}{\partial x}\hat{\imath} + \frac{\partial\varphi}{\partial y}\hat{\jmath} + \frac{\partial\varphi}{\partial z}\hat{k}\right)$

$$\nabla\varphi = (ay^2 + 3cz^2x^2)\hat{\imath} + (2axy + bz)\hat{\jmath} + (by + 2czx^3)\hat{k}$$

At $(-1,1,2)$

$$\nabla\varphi = (ay^2 + 3cz^2x^2)\hat{\imath} + (2axy + bz)\hat{\jmath} + (by + 2czx^3)\hat{k}$$

$$\nabla\varphi = (a + 12c)\hat{\imath} + (-2a + 2b)\hat{\jmath} + (b - 4c)\hat{k}$$

Since the D.D of φ has a maximum magnitude of 32units in the direction parallel to $y - axis$ is

$$\nabla\varphi.\hat{\jmath} = 32$$

$$\left[(a + 12c)\hat{\imath} + (-2a + 2b)\hat{\jmath} + (b - 4c)\hat{k}\right].\left[0\hat{\imath} + 1.\hat{\jmath} + 0\hat{k}\right] = 32$$

$\Rightarrow a + 12c = 0$ $\quad$ $-2a + 2b = 32$ $\quad$ $b - 4c = 0$

$\Rightarrow \boldsymbol{a = -12c}$ $\quad$ $\Rightarrow b - a = 16$ $\quad$ $\Rightarrow \boldsymbol{b = 4c}$

$$\Rightarrow 4c-(-12c)=16$$

$$\Rightarrow 16c=16$$

$\Rightarrow \mathbf{a=-12}$ $\quad\quad\quad\quad$ $\Rightarrow \boldsymbol{c=1}$ $\quad\quad\quad\quad$ $\Rightarrow \boldsymbol{b=4}$

$\therefore\ \boldsymbol{a=-12,\ b=4,\ c=1}$

8. Find the angle between the surfaces or normal surfaces $x^2+y^2+z^2=9$ and $z=x^2+y^2-3$ at $(2,-1,2)$.

Sol: Given $\varphi=x^2+y^2+z^2-9$ $\quad\quad$ $\psi=x^2+y^2-z-3$

$\frac{\partial\varphi}{\partial x}=2x$ $\quad\quad$ $\frac{\partial\psi}{\partial x}=2x$

$\frac{\partial\varphi}{\partial y}=2y$ $\quad\quad$ $\frac{\partial\psi}{\partial y}=2y$

$\frac{\partial\varphi}{\partial z}=2z$ $\quad\quad$ $\frac{\partial\psi}{\partial z}=-1$

Wkt, $\nabla\varphi=\frac{\partial\varphi}{\partial x}\hat{\imath}+\frac{\partial\varphi}{\partial y}\hat{\jmath}+\frac{\partial\varphi}{\partial z}\hat{k}$

$\nabla\varphi=(2x)\hat{\imath}+(2y)\hat{\jmath}+(2z)\hat{k}$

At $(2,-1,2)$

$$\nabla\varphi=(2.2)\hat{\imath}+\big(2.(-1)\big)\hat{\jmath}+(2.2)\hat{k}$$

$$\boldsymbol{\nabla\varphi=4\hat{\imath}-2\hat{\jmath}+4\hat{k}}$$

$$|\nabla\varphi|=\sqrt{4^2+(-2)^2+4^2}=\sqrt{36}$$

$$\boldsymbol{|\nabla\varphi|=6}$$

Also $\nabla\psi=\frac{\partial\psi}{\partial x}\hat{\imath}+\frac{\partial\psi}{\partial y}\hat{\jmath}+\frac{\partial\psi}{\partial z}\hat{k}$

$\nabla\psi=(2x)\hat{\imath}+(2y)\hat{\jmath}+(-1)\hat{k}$

At $(2,-1,2)$

$\nabla\psi=(2.2)\hat{\imath}+(2.(-1))\hat{\jmath}+(-1)\hat{k}$

$\boldsymbol{\nabla\psi=4\hat{\imath}-2\hat{\jmath}-\hat{k}}$

$$|\nabla\psi|=\sqrt{4^2+(-2)^2+(-1)^2}$$

$$|\nabla\psi|=\sqrt{21}$$

Angle between surface,

$$cos\theta = \widehat{n_1}.\widehat{n_2}$$

$$cos\theta = \frac{\nabla\varphi}{|\nabla\varphi|}.\frac{\nabla\psi}{|\nabla\psi|}$$

$$cos\theta = \frac{(\mathbf{4\hat{\imath} - 2\hat{\jmath} + 4\hat{k}})}{6}.\frac{(\mathbf{4\hat{\imath} - 2\hat{\jmath} - \hat{k}})}{\sqrt{21}}$$

$$cos\theta = \frac{16 + 4 - 4}{6\sqrt{21}} = \frac{16}{6\sqrt{21}} \qquad cos\theta = \frac{8}{3\sqrt{21}}$$

$$\boldsymbol{\theta = cos^{-1}\left(\frac{8}{3\sqrt{21}}\right)}$$

9. Find the angle between the normal to the surface $xy = z^2$ at the point $(4,1,2)$ & $(3,3,-3)$.

Sol: Given $\varphi = xy - z^2$

$\frac{\partial\varphi}{\partial x} = 1.y - 0$ $\qquad$ $\frac{\partial\varphi}{\partial y} = x.1 - 0$ $\qquad$ $\frac{\partial\varphi}{\partial z} = 0 - 2z$

$\frac{\partial\varphi}{\partial x} = y$ $\qquad$ $\frac{\partial\varphi}{\partial y} = x$ $\qquad$ $\frac{\partial\varphi}{\partial z} = -2z$

Wkt, $\nabla\varphi = \frac{\partial\varphi}{\partial x}\hat{\imath} + \frac{\partial\varphi}{\partial y}\hat{\jmath} + \frac{\partial\varphi}{\partial z}\hat{k}$

$$\nabla\varphi = (y)\hat{\imath} + (x)\hat{\jmath} + (-2z)\hat{k}$$

At $(4,1,2)$ $\qquad$ At $(3,3,-3)$

$\nabla\varphi_1 = (1)\hat{\imath} + (4)\hat{\jmath} + (-2.2)\hat{k}$ $\qquad$ $\nabla\varphi_2 = (3)\hat{\imath} + (3)\hat{\jmath} + (2.3)\hat{k}$

$\boldsymbol{\nabla\varphi_1 = \hat{\imath} + 4\hat{\jmath} - 4\hat{k}}$ $\qquad$ $\boldsymbol{\nabla\varphi_2 = 3\hat{\imath} + 3\hat{\jmath} + 6\hat{k}}$

$|\nabla\varphi_1| = \sqrt{1^2 + 4^2 + (-4)^2}$ $\qquad$ $|\nabla\varphi_2| = \sqrt{3^2 + 3^2 + 6^2} = \sqrt{54}$

$\boldsymbol{|\nabla\varphi_1| = \sqrt{33}}$ $\qquad$ $\boldsymbol{|\nabla\varphi_1| = 3\sqrt{6}}$

Angle between surface,

$$cos\theta = \widehat{n_1}.\widehat{n_2}$$

$$cos\theta = \frac{\nabla\varphi_1}{|\nabla\varphi_1|}.\frac{\nabla\varphi_2}{|\nabla\varphi_2|}$$

$$cos\theta = \frac{(\mathbf{\hat{\imath} + 4\hat{\jmath} - 4\hat{k}})}{\sqrt{33}}.\frac{(\mathbf{3\hat{\imath} + 3\hat{\jmath} + 6\hat{k}})}{3\sqrt{6}}$$

$$cos\theta = \frac{3+12-24}{3\sqrt{33.6}} = \frac{-9}{3.3\sqrt{22}}$$

$cos\theta = \frac{-1}{\sqrt{22}}$; $\boldsymbol{\theta = cos^{-1}\left[\frac{-1}{\sqrt{22}}\right]}$

<u>Divergence of a vector function</u>

The divergence of a vector function $\vec{F} = f_1\hat{\imath} + f_2\hat{\jmath} + f_3\hat{k}$, where f_1, f_2, f_3 are functions of x, y, z.It is denoted by $div\vec{F}$ and is defined as

$$div\vec{F} = \nabla.\vec{F}$$

$$div\vec{F} = \left(\frac{\partial}{\partial x}\hat{\imath} + \frac{\partial}{\partial y}\hat{\jmath} + \frac{\partial}{\partial z}\hat{k}\right).\left(f_1\hat{\imath} + f_2\hat{\jmath} + f_3\hat{k}\right)$$

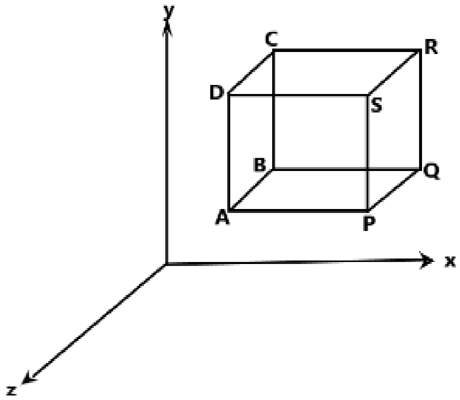
$$\boldsymbol{div\vec{F} = \nabla.\vec{F} = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z}}$$

Clearly $\boldsymbol{div\vec{F}}$ is a scalar quantity.

<u>Physical interpretation of Divergence</u>

Let us consider the motion of the fluid. Consider a small rectangular parallelepiped with edges $\delta_x, \delta_y, \delta_z$ parallel to the axes in the mass of fluid.

Let $\vec{V} = V_x\hat{\imath} + V_y\hat{\jmath} + V_z\hat{k}$ be the velocity of the fluid at (x, y, z).



Amount of the fluid flowing in through the face ABCD per unit time

$$= Velocity\ X\ Area\ of\ the\ face =\ V_x\ \delta y\delta z$$

Amount of the fluid flowing out through the face PQRS per unit time

$$= \left[V_x + \frac{\partial V_x}{\partial x}\ \delta x\right]\delta y\delta z$$

∴ The net decrease in the amount of fluid across these two faces is

$$= \left[V_x + \frac{\partial V_x}{\partial x}\ \delta x\right]\delta y\delta z - V_x\delta y\delta z$$

$$= \left[V_x + \frac{\partial V_x}{\partial x}\ \delta x - V_x\right]\delta y\delta z = \boldsymbol{\frac{\partial V_x}{\partial x}\ \delta x\delta y\delta z}$$

Similarly, the decrease in amount of fluid due to flow along the y − axis = $\frac{\partial V_y}{\partial y}\ \delta x\delta y\delta z$

The decrease in amount of fluid due to flow along the z − axis = $\frac{\partial V_z}{\partial z}\ \delta x\delta y\delta z$

Total decrease in amount of fluid inside the parallelepiped per unit time

$$=\left[\frac{\partial V_x}{\partial x}+\frac{\partial V_y}{\partial y}+\frac{\partial V_z}{\partial z}\right]\ \delta x\delta y\delta z$$

Hence the ratio of loss of fluid per unit volume = $\frac{\partial V_x}{\partial x}+\frac{\partial V_y}{\partial y}+\frac{\partial V_z}{\partial z}$

$$=\left[\frac{\partial}{\partial x}\hat{\imath}+\frac{\partial}{\partial y}\hat{\jmath}+\frac{\partial}{\partial z}\hat{k}\right].\left[V_x\hat{\imath}+V_y\hat{\jmath}+V_z\hat{k}\right]$$

$$=\nabla.\vec{V}$$

$$=div\vec{V}$$

Hence $div\vec{V}$ gives the rate of outflow per unit volume at a point of the fluid. If $\boldsymbol{div\vec{V}=0}$ everywhere in some region of space, then $\vec{V}$ is called ***Solenoidal Vector function*** and the fluids said to be *incompressible* i.e., there is no gain or loss in the volume element.

Curl of a vector function

The curl of a vector function $\vec{F}=f_1\hat{\imath}+f_2\hat{\jmath}+f_3\hat{k}$ is denoted by $curl\vec{F}$ and is defined as $\boldsymbol{curl\vec{F}}=\boldsymbol{\nabla X\vec{F}}$

$$curl\vec{F}=\begin{vmatrix}\hat{\imath} & \hat{\jmath} & \hat{k}\\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z}\\ f_1 & f_2 & f_3\end{vmatrix}$$

$$\boldsymbol{curl\vec{F}}=\left[\frac{\partial f_3}{\partial y}-\frac{\partial f_2}{\partial z}\right]\hat{\boldsymbol{\imath}}-\left[\frac{\partial f_3}{\partial x}-\frac{\partial f_1}{\partial z}\right]\hat{\boldsymbol{\jmath}}+\left[\frac{\partial f_2}{\partial x}-\frac{\partial f_1}{\partial y}\right]\hat{\boldsymbol{k}}$$

Clearly $\boldsymbol{curl\vec{F}}$ is a vector quantity.

Physical interpretation of Curl

Consider a rigid body rotating about a fixed axis through origin. Let the uniform angular velocity be $\vec{w}=w_1\hat{\imath}+w_2\hat{\jmath}+w_3\hat{k}$, w_1, w_2, w_3 are constants. The velocity $\vec{V}$ of any point P(x, y, z) on the body is given by $\vec{V}=\vec{w}\ X\ \vec{r}$, where $\vec{r}$ is the position vector of P.

Let $\vec{w}=w_1\hat{\imath}+w_2\hat{\jmath}+w_3\hat{k}$ and $\vec{r}=x\hat{\imath}+y\hat{\jmath}+z\hat{k}$

Consider $\vec{V}=\vec{w}\ X\ \vec{r}$

$$=\begin{vmatrix}\hat{\imath} & \hat{\jmath} & \hat{k}\\ w_1 & w_2 & w_3\\ x & y & z\end{vmatrix}$$

$$\vec{V}=(w_2z-w_3y)\hat{\imath}-(w_1z-w_3x)\hat{\jmath}+(w_1y-w_2x)\hat{k}$$

$$curl\vec{V}=\nabla X\vec{V}$$

$$curl\vec{V}=\nabla\ X\left[(w_2z-w_3y)\hat{\imath}-(w_1z-w_3x)\hat{\jmath}+(w_1y-w_2x)\hat{k}\right]$$

$$curl\vec{V} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ (w_2z - w_3y) & (w_3x - w_1z) & (w_1y - w_2x) \end{vmatrix}$$

$$curl\vec{V} = [w_1 - (-w_1)\,]\hat{\imath} - [-w_2 - w_2]\hat{\jmath} + [w_3 - (-w_3)]\hat{k}$$

$$curl\vec{V} = [2w_1\,]\hat{\imath} + [2w_2]\hat{\jmath} + [2w_3]\hat{k}$$

$$curl\vec{V} = 2\left[w_1\hat{\imath} + w_2\hat{\jmath} + w_3\hat{k}\right]$$

$$curl\vec{V} = 2\vec{w}$$

$$\boldsymbol{\vec{w}} = \frac{\mathbf{1}}{\mathbf{2}}\mathbf{curl}\boldsymbol{\vec{V}}$$

Thus the angular velocity of rotation at any point is equal to half of the curl of the velocity.

Note:

1. If $\boldsymbol{div\vec{F} = 0}$, then we say that $\vec{F}$ is **Solenoidal** vector.
2. If $\boldsymbol{curl\vec{F} = 0}$, then we say that $\vec{F}$ is **irrotational** vector.
3. Irrotational vector field is called as conservative field or potential field.
4. When $\vec{F}$ is irrotational there always exist a scalar point function such that $\boldsymbol{\nabla\varphi = \vec{F}}$, then φ is called a scalar potential of vector $\vec{F}$.

Problems

1. If $\vec{F} = xyz\hat{\imath} + 3x^2y\hat{\jmath} + (xz^2 - y^2z)\hat{k}$ **,** find the $div\vec{F}$ and $curl\vec{F}$ at $(2,-1,1)$.

Sol: Given $\vec{F} = xyz\hat{\imath} + 3x^2y\hat{\jmath} + (xz^2 - y^2z)\hat{k}$

Wkt $div\vec{F} = \nabla.\vec{F}$

$$div\vec{F} = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z}$$

$$div\vec{F} = \frac{\partial}{\partial x}(xyz) + \frac{\partial}{\partial y}(3x^2y) + \frac{\partial}{\partial z}(xz^2 - y^2z)$$

$$div\vec{F} = yz.1 + 3x^2.1 + (x.2z - y^2.1)$$

$$div\vec{F} = yz + 3x^2 + 2xz - y^2$$

At $(2,-1,1)$

$$div\vec{F} = (-1)(1) + 3.(2)^2 + 2(2)(1) - (-1)^2$$

$$div\vec{F} = -1 + 12 + 4 - 1$$

$\boldsymbol{div\vec{F}}$ **=14**

Also, $curl\vec{F} = \nabla\, X\, \vec{F}$

$$= \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ (xyz) & (3x^2y) & (xz^2 - y^2z) \end{vmatrix}$$

$$curl\vec{F} = \left[\frac{\partial}{\partial y}(xz^2 - y^2z) - \frac{\partial}{\partial z}(3x^2y)\right]\hat{\imath} - \left[\frac{\partial}{\partial x}(xz^2 - y^2z) - \frac{\partial}{\partial z}(xyz)\right]\hat{\jmath} + \left[\frac{\partial}{\partial x}(3x^2y) - \frac{\partial}{\partial y}(xyz)\right]\hat{k}$$

$$curl\vec{F} = [(-2yz) - 0]\hat{\imath} - [(z^2 - 0) - (xy)]\hat{\jmath} + [(6xy) - (xz)]\hat{k}$$

$$curl\vec{F} = [-2yz]\hat{\imath} - [z^2 - xy]\hat{\jmath} + [6xy - xz]\hat{k}$$

At (2,-1,1)

$$curl\vec{F} = [-2(-1)(1)]\hat{\imath} - [1^2 - 2(-1)]\hat{\jmath} + [6(2)(-1) - (2)(1)]\hat{k}$$

$$\boldsymbol{curl\vec{F} = 2\hat{\imath} - 3\hat{\jmath} - 14\hat{k}}$$

2. Find $div\vec{F}$ and $curl\vec{F}$ where $\vec{F} = grad(x^3 + y^3 + z^3 - 3xyz)$.

Sol: Let $\varphi = x^3 + y^3 + z^3 - 3xyz$

$$\vec{F} = grad\varphi$$

$$\vec{F} = \nabla\varphi$$

$$\vec{F} = \frac{\partial\varphi}{\partial x}\hat{\imath} + \frac{\partial\varphi}{\partial y}\hat{\jmath} + \frac{\partial\varphi}{\partial z}\hat{k}$$

$$\boldsymbol{\vec{F} = (3x^2 - 3yz)\hat{\imath} + (3y^2 - 3xz)\hat{\jmath} + (3z^2 - 3xy)\hat{k}}$$

Now, $div\vec{F} = \nabla.\vec{F}$

$$div\vec{F} = \frac{\partial}{\partial x}(3x^2 - 3yz) + \frac{\partial}{\partial y}(3y^2 - 3xz) + \frac{\partial}{\partial z}(3z^2 - 3xy)$$

$$div\vec{F} = (6x - 0) + (6y - 0) + (6z - 0)$$

$$\boldsymbol{div\vec{F} = 6x + 6y + 6z}$$

$$curl\vec{F} = \nabla\, X\, \vec{F}$$

$$curl\vec{F} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ (3x^2 - 3yz) & (3y^2 - 3xz) & (3z^2 - 3xy) \end{vmatrix}$$

$$curl\vec{F} = \left[\frac{\partial}{\partial y}(3z^2 - 3xy) - \frac{\partial}{\partial z}(3y^2 - 3xz)\right]\hat{\imath} - \left[\frac{\partial}{\partial x}(3z^2 - 3xy) - \frac{\partial}{\partial z}(3x^2 - 3yz)\right]\hat{\jmath} + \left[\frac{\partial}{\partial x}(3y^2 - 3xz) - \frac{\partial}{\partial y}(3x^2 - 3yz)\right]\hat{k}$$

$$curl\vec{F} = [(0 - 3x) - (0 - 3x)]\hat{\imath} - [(0 - 3y) - (0 - 3y)]\hat{\jmath} + [(0 - 3z) - (0 - 3z)]\hat{k}$$

$$curl\vec{F} = [-3x + 3x]\hat{\imath} - [-3y + 3y]\hat{\jmath} + [-3z + 3z]\hat{k}$$

$$curl\vec{F} = 0\hat{\imath} - 0\hat{\jmath} + 0\hat{k}$$

$$\boldsymbol{curl\vec{F} = \vec{0}}$$

3. If $\vec{F} = \nabla(xy^3z^2)$, find $div\vec{F}$ and $curl\vec{F}$ at $(1,-1,1)$.

Sol: Let $\varphi = xy^3z^2$

$$\vec{F} = \nabla\varphi$$

$$\vec{F} = \frac{\partial\varphi}{\partial x}\hat{\imath} + \frac{\partial\varphi}{\partial y}\hat{\jmath} + \frac{\partial\varphi}{\partial z}\hat{k}$$

$$\boldsymbol{\vec{F} = (y^3z^2)\hat{\imath} + (3xy^2z^2)\hat{\jmath} + (2xy^3z)\hat{k}}$$

Now, $div\vec{F} = \nabla.\vec{F}$

$$div\vec{F} = \frac{\partial}{\partial x}(y^3z^2) + \frac{\partial}{\partial y}(3xy^2z^2) + \frac{\partial}{\partial z}(2xy^3z)$$

$$div\vec{F} = (0) + (6xyz^2) + (2xy^3)$$

$$div\vec{F} = 6xyz^2 + 2xy^3$$

At $(1,-1,1)$

$$div\vec{F} = 6(1)(-1)(1)^2 + 2(1)(-1)^3 = -6 - 2$$

$$\boldsymbol{div\vec{F} = -8}$$

Also, $curl\vec{F} = \nabla\, X\, \vec{F}$

$$curl\vec{F} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ (\boldsymbol{y^3z^2}) & (\boldsymbol{3xy^2z^2}) & (\boldsymbol{2xy^3z}) \end{vmatrix}$$

$$curl\vec{F} = \left[\frac{\partial}{\partial y}(\boldsymbol{2xy^3z}) - \frac{\partial}{\partial z}(\boldsymbol{3xy^2z^2})\right]\hat{\imath} - \left[\frac{\partial}{\partial x}(\boldsymbol{2xy^3z}) - \frac{\partial}{\partial z}(\boldsymbol{y^3z^2})\right]\hat{\jmath} + \left[\frac{\partial}{\partial x}(\boldsymbol{3xy^2z^2}) - \frac{\partial}{\partial y}(\boldsymbol{y^3z^2})\right]\hat{k}$$

$$curl\vec{F} = [(\boldsymbol{6xy^2z^2}) - (\boldsymbol{6xy^2z^2})]\hat{\imath} - [(\boldsymbol{2y^3z}) - (\boldsymbol{2y^3z})]\hat{\jmath} + [(\boldsymbol{3y^2z^2}) - (\boldsymbol{3y^2z^2})]\hat{k}$$

$$curl\vec{F} = 0\hat{\imath} - 0\hat{\jmath} + 0\hat{k}$$

$$\boldsymbol{curl\vec{F} = \vec{0}}$$

4. If $\vec{F} = (3x^2y - z)\hat{\imath} + (xz^3 + y^4)\hat{\jmath} - 2x^3z^2\hat{k}$, find $grad(div\vec{F})$ at $(2,-1,0)$.

Sol: Given $\vec{F} = (3x^2y - z)\hat{\imath} + (xz^3 + y^4)\hat{\jmath} - 2x^3z^2\hat{k}$

Wkt, $div\vec{F} = \nabla.\vec{F}$

$$div\vec{F} = \frac{\partial}{\partial x}(3x^2y - z) + \frac{\partial}{\partial y}(xz^3 + y^4) + \frac{\partial}{\partial z}(-2x^3z^2)$$

$$div\vec{F} = (6xy) + (0 + 4y^3) + (-4x^3z)$$

$$\boldsymbol{div\vec{F} = 6xy + 4y^3 - 4x^3z = \varphi(Say)}$$

$\frac{\partial\varphi}{\partial x} = 6y - 12x^2z \qquad \frac{\partial\varphi}{\partial y} = 6x + 12y^2 \qquad \frac{\partial\varphi}{\partial z} = -4x^3$

Now, $grad\left(div\vec{F}\right) = grad\varphi$

$$grad\left(div\vec{F}\right) = \nabla\varphi$$

$$= \frac{\partial\varphi}{\partial x}\hat{\imath} + \frac{\partial\varphi}{\partial y}\hat{\jmath} + \frac{\partial\varphi}{\partial z}\hat{k}$$

$$grad(div\vec{F}) \;= (6y - 12x^2z)\hat{\imath} + (6x + 12y^2)\hat{\jmath} + (-4x^3)\hat{k}$$

At $(2, -1, 0)$

$$grad(div\vec{F}) \;= (6(-1) - 0)\hat{\imath} + (6(2) + 12(-1)^2)\hat{\jmath} + (-4(2)^3)\hat{k}$$

$$\boldsymbol{grad\left(div\vec{F}\right) = \; 6\hat{\imath} + 24\hat{\jmath} - 32\hat{k}}$$

6. If $\vec{F} = x^2\hat{\imath} + xy\hat{\jmath} + xz\hat{k}$, find $curl(curl\vec{F})$.

Sol: Given $\vec{F} = x^2\hat{\imath} + xy\hat{\jmath} + xz\hat{k}$

$$curl\vec{F} = \nabla\, X\, \vec{F} curl\vec{F} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ \left(\boldsymbol{x^2}\right) & (\boldsymbol{xy}) & (\boldsymbol{xz}) \end{vmatrix}$$

$$curl\vec{F} = \left[\frac{\partial}{\partial y}(xz) - \frac{\partial}{\partial z}(xy)\right]\hat{\imath} - \left[\frac{\partial}{\partial x}(xz) - \frac{\partial}{\partial z}(x^2)\right]\hat{\jmath} + \left[\frac{\partial}{\partial x}(xy) - \frac{\partial}{\partial y}(x^2)\right]\hat{k}$$

$$curl\vec{F} = [(0) - (0)]\hat{\imath} - [(z) - (0)]\hat{\jmath} + [(y) - (0)]\hat{k}$$

$$\boldsymbol{curl\vec{F} = 0\hat{\imath} - z\hat{\jmath} + y\hat{k}}$$

Now,

$$curl(curl\vec{F}) = \nabla X(curl\vec{F})$$

$$curl\left(curl\vec{F}\right) = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ (\boldsymbol{0}) & (-\boldsymbol{z}) & (\boldsymbol{y}) \end{vmatrix}$$

$$curl(curl\vec{F}) = \left[\frac{\partial}{\partial y}(y) - \frac{\partial}{\partial z}(-z)\right]\hat{\imath} - \left[\frac{\partial}{\partial x}(y) - \frac{\partial}{\partial z}(0)\right]\hat{\jmath} + \left[\frac{\partial}{\partial x}(-z) - \frac{\partial}{\partial y}(0)\right]\hat{k}$$

$$curl(curl\vec{F}) = [(1) - (-1)]\hat{\imath} - [(0) - (0)]\hat{\jmath} + [(0) - (0)]\hat{k}$$

$$\boldsymbol{curl(curl\vec{F}) = 2\hat{\imath}}$$

7. Show that $\vec{F} = \frac{x\hat{\imath} + y\hat{\jmath}}{(x^2+y^2)}$ is both solenoidal and irrotational.

Sol: Given $\vec{F} = \frac{x}{(x^2+y^2)}\hat{\imath} + \frac{y}{(x^2+y^2)}\hat{\jmath}$

Wkt, $div\vec{F} = \nabla.\vec{F}$

$$div\vec{F} = \frac{\partial}{\partial x}\left(\frac{x}{(x^2+y^2)}\right)+\frac{\partial}{\partial y}\left(\frac{y}{(x^2+y^2)}\right)$$

$$div\vec{F} = \frac{(x^2+y^2).1-x.2x}{(x^2+y^2)^2} + \frac{(x^2+y^2).1-y.2y}{(x^2+y^2)^2}$$

$$div\vec{F} = \frac{x^2+y^2-2x^2+x^2+y^2-2y^2}{(x^2+y^2)^2}$$

$$\boldsymbol{div\vec{F} = 0}$$

$\therefore$ $\boldsymbol{\vec{F}}$ **is Solenoidal.**

Now, $curl\vec{F} = \nabla\, X\, \vec{F}$

$$curl\vec{F} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ \dfrac{x}{(x^2+y^2)} & \dfrac{y}{(x^2+y^2)} & 0 \end{vmatrix}$$

$$curl\vec{F} = \left[\frac{\partial}{\partial y}(0) - \frac{\partial}{\partial z}\left(\frac{y}{(x^2+y^2)}\right)\right]\hat{\imath} - \left[\frac{\partial}{\partial x}(0) - \frac{\partial}{\partial z}\left(\frac{x}{(x^2+y^2)}\right)\right]\hat{\jmath} + \left[\frac{\partial}{\partial x}\left(\frac{y}{(x^2+y^2)}\right) - \frac{\partial}{\partial y}\left(\frac{x}{(x^2+y^2)}\right)\right]\hat{k}$$

$$curl\vec{F} = [(0)-(0)]\hat{\imath} - [(0)-(0)]\hat{\jmath} + \left[\frac{-2xy}{(x^2+y^2)} + \frac{2xy}{(x^2+y^2)}\right]\hat{k}$$

$$curl\vec{F} = 0\hat{\imath} - 0\hat{\jmath} + 0\hat{k}$$

$$\boldsymbol{curl\vec{F} = \vec{0}}$$

$\therefore$ ***$\vec{F}$ is irrotational.***

8. P.T $\vec{F} = (y+z)\hat{\imath} + (z+x)\hat{\jmath} + (x+y)\hat{k}$ is irrotational. Also find a scalar point function φ such that $\vec{F} = \nabla\varphi$.

Sol: *Given* $\vec{F} = (y+z)\hat{\imath} + (z+x)\hat{\jmath} + (x+y)\hat{k}$

$$curl\vec{F} = \nabla\, X\, \vec{F}$$

$$curl\vec{F} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ (y+z) & (z+x) & (x+y) \end{vmatrix}$$

$$curl\vec{F} = \left[\frac{\partial}{\partial y}(x+y) - \frac{\partial}{\partial z}(z+x)\right]\hat{\imath} - \left[\frac{\partial}{\partial x}(x+y) - \frac{\partial}{\partial z}(y+z)\right]\hat{\jmath} + \left[\frac{\partial}{\partial x}(z+x) - \frac{\partial}{\partial y}(y+z)\right]\hat{k}$$

$$curl\vec{F} = [(1)-(1)]\hat{\imath} - [(1)-(1)]\hat{\jmath} + [(1)-(1)]\hat{k}$$

$$curl\vec{F} = 0\hat{\imath} - 0\hat{\jmath} + 0\hat{k}$$

$$\boldsymbol{curl\vec{F} = \vec{0}} \therefore\ \boldsymbol{\vec{F}}\ \textbf{\textit{is irrotational.}}$$

<u>To find φ</u>

Consider $\nabla\varphi = \vec{F}$

$$\frac{\partial\varphi}{\partial x}\hat{\imath} + \frac{\partial\varphi}{\partial y}\hat{\jmath} + \frac{\partial\varphi}{\partial z}\hat{k} = (y+z)\hat{\imath} + (z+x)\hat{\jmath} + (x+y)\hat{k}$$

$\frac{\partial\varphi}{\partial x} = y + z$ $\quad\quad$ $\frac{\partial\varphi}{\partial y} = z + x$ $\quad\quad$ $\frac{\partial\varphi}{\partial z} = x + y$

Integrating we get

$\varphi = (y+z)\int 1\,dx$ $\quad\quad$ $\varphi = (z+x)\int 1\,dy$ $\quad\quad$ $\varphi = (x+y)\int 1\,dz$

$\varphi = (y+z)x + f(y,z)$ $\quad\quad$ $\varphi = (z+x)y + f(x,z)$ $\quad\quad$ $\varphi = (x+y)z + f(x,y)$

$\varphi = xy + xz + f(y,z)$ $\quad\quad$ $\varphi = yz + xy + f(x,z)$ $\quad\quad$ $\varphi = xz + yz + f(x,y)$

$\therefore$ $\boldsymbol{\varphi = xy + xz + yz}$ **,** *where* $f(y,z) = yz$ **,** $f(x,z) = xz$ **,** $f(x,y) = xy$

9. If $\vec{F} = (axy + z^3)\hat{\imath} + (3x^2 - z)\hat{\jmath} + (bxz^2 - y)\hat{k}$ *; if* $\vec{F}$ *is irrotational find constants a and b. Also find scalar function* φ *such that* $\vec{F} = \nabla\varphi$.

Sol: *Given* $\vec{F} = (axy + z^3)\hat{\imath} + (3x^2 - z)\hat{\jmath} + (bxz^2 - y)\hat{k}$

$$curl\vec{F} = \nabla\, X\, \vec{F}$$

$$curl\vec{F} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ (axy + z^3) & (3x^2 - z) & (bxz^2 - y) \end{vmatrix}$$

$$curl\vec{F} = \left[\frac{\partial}{\partial y}(bxz^2 - y) - \frac{\partial}{\partial z}(3x^2 - z)\right]\hat{\imath} - \left[\frac{\partial}{\partial x}(bxz^2 - y) - \frac{\partial}{\partial z}(axy + z^3)\right]\hat{\jmath} + \left[\frac{\partial}{\partial x}(3x^2 - z) - \frac{\partial}{\partial y}(axy + z^3)\right]$$

$$curl\vec{F} = [(-1) - (-1)]\hat{\imath} - [(bz^2) - (3z^2)]\hat{\jmath} + [(6x) - (ax)]\hat{k}curl\vec{F}$$
$$= 0\hat{\imath} - (bz^2 - 3z^2)\hat{\jmath} + (6x - ax)\hat{k}$$

Since $\vec{F}$ *is irrotational,*

$curl\vec{F} = \vec{0}$

$$0\hat{\imath} - (bz^2 - 3z^2)\hat{\jmath} + (6x - ax)\hat{k} = 0\hat{\imath} - 0\hat{\jmath} + 0\hat{k}$$

$bz^2 - 3z^2 = 0$ $\quad\quad$ $6x - ax = 0$

$z^2(b - 3) = 0$ $\quad\quad$ $x(6 - a) = 0$

$b - 3 = 0$ $\quad\quad$ $6 - a = 0$

$\boldsymbol{b = 3}$ $\quad\quad$ $\boldsymbol{a = 6}$

Thus, $\boldsymbol{\vec{F} = (6xy + z^3)\hat{\imath} + (3x^2 - z)\hat{\jmath} + (3xz^2 - y)\hat{k}}$

<u>To find φ</u>

Consider $\nabla\varphi = \vec{F}$

$$\frac{\partial\varphi}{\partial x}\hat{\imath} + \frac{\partial\varphi}{\partial y}\hat{\jmath} + \frac{\partial\varphi}{\partial z}\hat{k} = (6xy + z^3)\hat{\imath} + (3x^2 - z)\hat{\jmath} + (3xz^2 - y)\hat{k}$$

$\frac{\partial\varphi}{\partial x} = 6xy + z^3$ $\qquad$ $\frac{\partial\varphi}{\partial y} = 3x^2 - z$ $\qquad$ $\frac{\partial\varphi}{\partial z} = 3xz^2 - y$

Integrating we get

$\varphi = 6y\int x\,dx + z^3\int 1\,dx$ $\qquad$ $\varphi = 3x^2\int 1\,dy - z\int 1\,dy$ $\qquad$ $\varphi = 3x\int z^2dz - y\int 1\,dz$

$\varphi = 6y\left(\frac{x^2}{2}\right) + z^3x + f(y,z)$ $\qquad$ $\varphi = 3x^2y - zy + f(x,z)$ $\qquad$ $\varphi = 3x\left(\frac{z^3}{3}\right) - yz + f(x,y)$

$\varphi = 3x^2y + z^3x + f(y,z)$ $\qquad$ $\varphi = 3x^2y - zy + f(x,z)$ $\qquad$ $\varphi = xz^3 - yz + f(x,y)$

$\therefore$ $\boldsymbol{\varphi = 3x^2y + xz^3 - yz}$ **,** *where* $f(y,z) = -yz$ **,** $f(x,z) = xz^3$**,** $f(x,y) = 3x^2y$

VECTOR INTEGRATION

Line Integral

Consider a curve C *in space which consists of infinitesimally small elements of length* dr. *Then the line integral of a vector* $\vec{A}(x,y,z)$ *along the curve* C *is defined to be the sum of the scalar products of* $\vec{A}$ *,*$d\vec{r}$ *and is represented by* $\int_C \vec{A}\,.\,d\vec{r}$.

If $\vec{F}$ *is the force acted upon by a particle in displacing it along the curve* C *then* $\int_C \vec{F}\,.\,d\vec{r}$ *represents the total work done by a force , it also represents the circulation of* $\vec{F}$ *about* C *where* $\vec{F}$ *represents the velocity of a fluid.*

$\vec{F}$ *is said to be irrotational if* $\oint_C \vec{F}.d\vec{r} = 0$.

<u>Problems</u>

1. If $\vec{F} = (5xy - 6x^2)\hat{\imath} + (2y - 4x)\hat{\jmath}$, *evaluate* $\int_C \vec{F}\,.\,d\vec{r}$ *where* C *is the curve* $y = x^3$ *from the point* $(1,1)$ *to the point* $(2,8)$.

Sol: *Given* $\vec{F} = (5xy - 6x^2)\hat{\imath} + (2y - 4x)\hat{\jmath}$

$$d\vec{r} = dx\hat{\imath} + dy\hat{\jmath} + dz\hat{k}$$

Consider, $\vec{F}.d\vec{r} = [(5xy - 6x^2)\hat{\imath} + (2y - 4x)\hat{\jmath}].\left[dx\hat{\imath} + dy\hat{\jmath} + dz\hat{k}\right]$

$$\vec{F}.d\vec{r} = (5xy - 6x^2)dx + (2y - 4x)dy \quad \text{---} \quad \textbf{(1)}$$

In C : $y = x^3$ $\qquad$ *Points:* $(1,1)$ $\quad$ $(2,8)$

$dy = 3x^2dx$

(1)$\Rightarrow$ $\vec{F}.d\vec{r} = (5x^4 - 6x^2)dx + (2x^3 - 4x).3x^2dx$

$$\vec{F}.d\vec{r} = (5x^4 - 6x^2 + 6x^5 - 12x^3)dx$$

$$\int_C \vec{F}\,.\,d\vec{r} = \int_1^2 (5x^4 - 6x^2 + 6x^5 - 12x^3)dx$$

$$\int_C \vec{F}\,.\,d\vec{r} = 5\left[\frac{x^5}{5}\right]_{x=1}^{x=2} - 6\left[\frac{x^3}{3}\right]_{x=1}^{x=2} + 6\left[\frac{x^6}{6}\right]_{x=1}^{x=2} - 12\left[\frac{x^4}{4}\right]_{x=1}^{x=2}$$

$= (2^5 - 1) - 2(2^3 - 1) + (2^6 - 1) - 3(2^4 - 1)$

$= 31 - 14 + 63 - 45$

$$\int_C \boldsymbol{\vec{F}\,.\,d\vec{r} = 35}$$

2. Evaluate $\int_C \vec{F}\,.\,d\vec{r}$ *along the circle* $x^2 + y^2 = 4$ *, where* $\vec{F} = 3xy\hat{\imath} - y\hat{\jmath} + 2z\hat{k}$.

Sol: *Given* $\vec{F} = 3xy\hat{\imath} - y\hat{\jmath} + 2z\hat{k}$

$d\vec{r} = dx\hat{\imath} + dy\hat{\jmath} + dz\hat{k}$

Consider, $\vec{F}.d\vec{r} = [3xy\hat{\imath} - y\hat{\jmath} + 2z\hat{k}].[dx\hat{\imath} + dy\hat{\jmath} + dz\hat{k}]$

$\vec{F}.d\vec{r} = (3xy)dx + (-y)dy + (2z)dz$ --- ***(1)***

In C : $x^2 + y^2 = 4$, $z = 0$

$x^2 + y^2 = 2^2$ $(x^2 + y^2 = r^2)$

Put $x = rcos\theta$; $y = rsin\theta$; $z = 0$

$x = 2cos\theta$; $y = 2sin\theta$

$dx = -2sin\theta d\theta$; $dy = 2cos\theta d\theta$; $dz = 0$

$\theta : \theta = 0$ *to* $\theta = 2\pi$

(1)$\Rightarrow \vec{F}.d\vec{r} = (3.\ 2cos\theta.\ 2sin\theta)(-2sin\theta d\theta) + (-2sin\theta).(2cos\theta d\theta)$

$\vec{F}.d\vec{r} = (-24cos\theta sin^2\theta - 4sin\theta cos\theta)d\theta$

$$\int_C \vec{F}\,.\,d\vec{r} = \int_{\theta=0}^{2\pi} (-24sin^2\theta cos\theta - 4sin\theta cos\theta)d\theta$$

$\int_C \vec{F}\,.\,d\vec{r} = -24\left[\frac{sin^3\theta}{3}\right]_0^{2\pi} - 4\left[\frac{sin^2\theta}{2}\right]_0^{2\pi}$ $\left\{\int [f(x)]^n f'(x)dx = \frac{[f(x)]^{n+1}}{n+1}\right\}$

$$\int_C \vec{F}\,.\,d\vec{r} = -8(0-0) - 2(0-0) \qquad \{sin2\pi = 0 = sin0\}$$

$$\int_C \vec{F}\,.\,d\vec{r} = 0.$$

3. If $\vec{F} = (3x^2 + 6y)\hat{\imath} - (14yz)\hat{\jmath} + (20xz^2)\hat{k}$*, evaluate* $\int_C \vec{F}\,.\,d\vec{r}$ *from the point* $(0,0,0)$ *to*

$(1,1,1)$ *along the curve given by* $x = t,\ y = t^2\ ,\ z = t^3$.

Sol: *Given* $\vec{F} = (3x^2 + 6y)\hat{\imath} - (14yz)\hat{\jmath} + (20xz^2)\hat{k}$

$$d\vec{r} = dx\hat{\imath} + dy\hat{\jmath} + dz\hat{k}$$

Consider, $\vec{F}.d\vec{r} = \left[(3x^2 + 6y)\hat{\imath} - (14yz)\hat{\jmath} + (20xz^2)\hat{k}\right].\left[dx\hat{\imath} + dy\hat{\jmath} + dz\hat{k}\right]$

$$\vec{F}.d\vec{r} = (3x^2 + 6y)dx + (-14yz)dy + (20xz^2)dz \quad \text{---} \quad \textbf{(1)}$$

In C : $x = t \quad ; \quad y = t^2 \quad ; \quad z = t^3$

$dx = dt \quad ; \quad dy = 2tdt \quad ; \quad dz = 3t^2dt$

$t: t = 0 \quad to \quad t = 1$

(1)$\Rightarrow \vec{F}.d\vec{r} = (3t^2 + 6t^2)dt - (14t^2t^3).2tdt + (20.t.t^6)3t^2dt$

$\vec{F}.d\vec{r} = (9t^2 - 28t^6 + 60t^9)dt$

$$\int_C \vec{F}\,.d\vec{r} = \int_0^1 (9t^2 - 28t^6 + 60t^9)dt$$

$$\int_C \vec{F}\,.d\vec{r} = 9\left[\frac{t^3}{3}\right]_{t=0}^{t=1} - 28\left[\frac{t^7}{7}\right]_{t=0}^{t=1} + 60\left[\frac{t^{10}}{10}\right]_{t=0}^{t=1}$$

$= 3(1-0) - 4(1-0) + 6(1-0)$

$= 3 - 4 + 6$

$$\int_C \vec{\boldsymbol{F}}\,.\boldsymbol{d}\vec{\boldsymbol{r}} = \mathbf{5}$$

4. *If* $\vec{F} = (x^2)\hat{\imath} + (xy)\hat{\jmath}$, *evaluate* $\int_C \vec{F}\,.d\vec{r}$ *from* $(0,0)$ *to* $(1,1)$ *along i) the line* $y = x$

ii) the parabola $y = \sqrt{x}$.

Sol: *Given* $\vec{F} = x^2\hat{\imath} + xy\hat{\jmath}$

$$d\vec{r} = dx\hat{\imath} + dy\hat{\jmath} + dz\hat{k}$$

Consider, $\vec{F}.d\vec{r} = [(x^2)\hat{\imath} + (xy)\hat{\jmath}].\left[dx\hat{\imath} + dy\hat{\jmath} + dz\hat{k}\right]$

$$\vec{F}.d\vec{r} = (x^2)dx + (xy)dy \quad \text{---} \quad \textbf{(1)}$$

i) *Along* $y = x$ **Point:** $(0,0)$ *to* $(1,1)$

$dy = dx$

(1)$\Rightarrow \vec{F}.d\vec{r} = (x^2)dx + (x.x)dx$

$\vec{F}.d\vec{r} = (x^2 + x^2)dx$

$$\vec{F}.d\vec{r} = 2x^2dx$$

$$\int_C \vec{F}\,.d\vec{r} = \int_0^1 2x^2dx$$

$$\int_C \vec{F}\,.\,d\vec{r} = 2\left[\frac{x^3}{3}\right]_{x=0}^{x=1}$$

$$\int_C \vec{F}\,.\,d\vec{r} = \frac{2}{3}(1-0)$$

$$\int_C \boldsymbol{\vec{F}\,.\,d\vec{r}} = \mathbf{\frac{2}{3}}$$

ii) Along $y=\sqrt{x}$ $\quad ; y^2 = x$ **Point:** $(0,0)$ *to* $(1,1)$

$2ydy = dx$

(1)$\Rightarrow \vec{F}.d\vec{r} = (y^4)2ydy + (y^2.y)dy$

$\vec{F}.d\vec{r} = (2y^5 + y^3)dy$

$$\int_C \vec{F}\,.\,d\vec{r} = \int_0^1 (2y^5 + y^3)dy$$

$$\int_C \vec{F}\,.\,d\vec{r} = 2\left[\frac{y^6}{6}\right]_{y=0}^{y=1} + \left[\frac{y^4}{4}\right]_{y=0}^{y=1}$$

$$\int_C \vec{F}\,.\,d\vec{r} = \frac{1}{3}(1-0) + \frac{1}{4}(1-0)$$

$$\int_C \vec{F}\,.\,d\vec{r} = \frac{1}{3} + \frac{1}{4}$$

$$\int_C \boldsymbol{\vec{F}\,.\,d\vec{r}} = \mathbf{\frac{7}{12}}$$

5. Find the total work done by a force $\vec{F} = 2xy\hat{\imath} - 4z\hat{\jmath} + 5x\hat{k}$ along the curve $x = t^2$, *y=(2t+1)* , $z = t^3$ from the point $t = 1$ to $t = 2$.

Sol: Given $\vec{F} = 2xy\hat{\imath} - 4z\hat{\jmath} + 5x\hat{k}$

$d\vec{r} = dx\hat{\imath} + dy\hat{\jmath} + dz\hat{k}$

Consider, $\vec{F}.d\vec{r} = \left[2xy\hat{\imath} - 4z\hat{\jmath} + 5x\hat{k}\right].\left[dx\hat{\imath} + dy\hat{\jmath} + dz\hat{k}\right]$

$\vec{F}.d\vec{r} = 2xydx - 4zdy + 5xdz$ **--- (1)**

In $C : x = t^2$; $\quad y = 2t + 1 \quad ; \quad z = t^3$

$dx = 2t\mathrm{dt}$; $\quad dy = 2\mathrm{dt} \quad ; \quad dz = 3t^2dt$

$t\!: t = 1 \;\; to \;\; t = 2$

(1)$\Rightarrow \vec{F}.d\vec{r} = (2t^2(2t+1))2tdt - 4(t^3).2dt + (5t^2)3t^2dt$

$\vec{F}.d\vec{r} = (8t^4 + 4t^3 - 8t^3 + 15t^4)dt$

$$\vec{F}.d\vec{r} = (23t^4 - 4t^3)dt$$

The work done by a force $= \int_C \vec{F}.d\vec{r}$

$$\int_C \vec{F}.d\vec{r} = \int_{t=1}^{2} (23t^4 - 4t^3)dt$$

$= 23\left[\frac{t^5}{5}\right]_{t=1}^{2} - 4\left[\frac{t^4}{4}\right]_{t=1}^{2}$

$= \frac{23}{5}(2^5 - 1) - (2^4 - 1)$

$= \frac{23(31)}{5} - 15 = \frac{713-75}{5}$

Work done , $\int_C \boldsymbol{\vec{F}.d\vec{r}} = \frac{\mathbf{638}}{\mathbf{5}}$ **units**

6. Find the total work done by a force $\vec{F} = 3x^2\hat{\imath} + (2xz - y)\hat{\jmath} + z\hat{k}$ when it moves a particle from the point $t = 0$ and $t = 2$ along the curve $x = t,\ y = \frac{t^2}{4}, z = \frac{3t^3}{8}$.

Sol: Given $\vec{F} = 3x^2\hat{\imath} + (2xz - y)\hat{\jmath} + z\hat{k}$

$d\vec{r} = dx\hat{\imath} + dy\hat{\jmath} + dz\hat{k}$

Consider, $\vec{F}.d\vec{r} = \left[3x^2\hat{\imath} + (2xz - y)\hat{\jmath} + z\hat{k}\right].\left[dx\hat{\imath} + dy\hat{\jmath} + dz\hat{k}\right]$

$\vec{F}.d\vec{r} = 3x^2dx + (2xz - y)dy + zdz$ **--- (1)**

In $C : x = t$; $y = \frac{t^2}{4}$; $z = \frac{3t^3}{8}$

$dx = dt$; $dy = \frac{2t}{4}dt = \frac{t}{2}dt$; $dz = \frac{(3)3t^2}{8}dt = \frac{9t^2}{8}dt$

$t: t = 0 \quad to \quad t = 2$

(1)⇒ $\vec{F}.d\vec{r} = 3t^2dt + \left[2t\left(\frac{3t^3}{8}\right) - \frac{t^2}{4}\right]\frac{t}{2}dt + \frac{3t^3}{8}.\frac{9t^2}{8}dt$

$\vec{F}.d\vec{r} = \left(3t^2 + \frac{3t^5}{8} - \frac{t^3}{8} + \frac{27t^5}{64}\right)dt$

$$\vec{F}.d\vec{r} = \left(3t^2 - \frac{t^3}{8} + \frac{51t^5}{64}\right)dt$$

The work done by a force $= \int_C \vec{F}.d\vec{r}$

$$\int_C \vec{F}.d\vec{r} = \int_{t=0}^{2} \left(3t^2 - \frac{t^3}{8} + \frac{51t^5}{64}\right)dt$$

$= 3\left[\frac{t^3}{3}\right]_{t=0}^{2} - \frac{1}{8}\left[\frac{t^4}{4}\right]_{t=0}^{2} + \frac{51}{64}\left[\frac{t^6}{6}\right]_{t=0}^{2}$

$= (2^3 - 0) - \frac{1}{32}(2^4 - 0) + \frac{51}{384}(2^6 - 0)$

$= 8 - \frac{166}{32} + \frac{51}{384}(64)$

$$\int_C \vec{F}.d\vec{r} = 8 - \frac{1}{2} + \frac{51}{6}$$

$$\int_C \vec{F}.d\vec{r} = \frac{48 - 3 + 51}{6}$$

$$\int_C \vec{F}\,.\,d\vec{r} = \frac{96}{6}$$

Work done, $\boldsymbol{\int_C \vec{F}\,.\,d\vec{r} = 16\ units}$

Green's Theorem

Let $M(x, y)$ and $N(x, y)$ be two functions defined in region $'R'$ and the $xy - Plane$ with simple closed curve C has its boundary , then $\oint_C \boldsymbol{Mdx + Ndy} = \iint_R \left[\frac{\partial M}{\partial x} - \frac{\partial N}{\partial y}\right]\boldsymbol{dydx}$

Note:

1. Area $= \iint_R \boldsymbol{dydx} = \frac{1}{2}\int_C (\boldsymbol{xdy - ydx})$

Problems

1. Evaluate $\int_C \boldsymbol{(xy - x^2)dx + x^2ydy}$ where C is the closed curve bounded by $y = 0$, $x = 1$ and $y = x$.

Sol: *Green's Theorem:* $\oint_C \boldsymbol{Mdx + Ndy} = \iint_R \left[\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right]\boldsymbol{dydx}$ --- (1)

Given $\int_C (xy - x^2)dx + x^2ydy$

Here, $\mathrm{M} = xy - x^2$ $\quad N = x^2y$

$\boldsymbol{\frac{\partial M}{\partial y} = x}$ $\quad \boldsymbol{\frac{\partial N}{\partial x} = 2xy}$

In 'R'

$$x: x = 0\ \ to\ \ x = 1$$

$$y: y = 0\ \ to\ \ y = x$$

$(1) \Rightarrow \int_C (xy - x^2)dx + x^2ydy = \int_{x=0}^{1}\int_{y=0}^{x}(2xy - x)dydx$

$$= \int_{x=0}^{1}\left[2x\left[\frac{y^2}{2}\right]_0^x - x[y]_0^x\right]dx$$

$$= \int_{x=0}^{1}[x(x^2 - 0) - x(x - 0)]dx$$

$$\int_C (xy - x^2)dx + x^2ydy = \int_{x=0}^{1} (x^3 - x^2)\, dx$$

$$= \left[\frac{x^4}{4}\right]_{x=0}^{1} - \left[\frac{x^3}{3}\right]_{x=0}^{1}$$

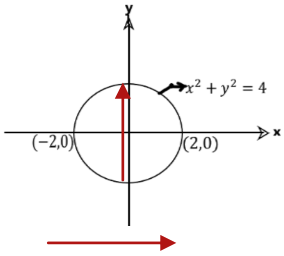
$$= \frac{1}{4} - \frac{1}{3} = \frac{3-4}{12}$$

$$\boldsymbol{\int_C (xy - x^2)dx + x^2ydy = \frac{-1}{12}}$$

2. Use Green's theorem to evaluate $\int_C \boldsymbol{(x^2 + y^2)dx + 3x^2ydy}$ where C is the circle $x^2 + y^2 = 4$ traced in the positive sign.

Sol: *Green's Theorem:* $\oint_C \boldsymbol{Mdx + Ndy = \iint_R \left[\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right] dydx}$ --- (1)

Given $\int_C (x^2 + y^2)dx + 3x^2ydy$



Here, $\text{M} = x^2 + y^2$ $\quad N = 3x^2$

$\boldsymbol{\frac{\partial M}{\partial y} = 2y}$ $\quad \boldsymbol{\frac{\partial N}{\partial x} = 6xy}$

<u>In 'R'</u>

$$x: x = -2 \ \ to \ \ x = 2$$

$$y: y = -\sqrt{4 - x^2} \quad to \ \ y = \sqrt{4 - x^2}$$

(1)⇒ $\int_C (x^2 + y^2)dx + 3x^2ydy = \int_{x=-2}^{2} \int_{y=-\sqrt{4-x^2}}^{\sqrt{4-x^2}} (6xy - 2y)dydx$

$$= \int_{x=-2}^{2} \left[6x \left[\frac{y^2}{2}\right]_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} - 2\left[\frac{y^2}{2}\right]_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}}\right] dx$$

$$= \int_{x=-2}^{2} \left[3x\left[\left[\sqrt{4 - x^2}\right]^2 - \left[-\sqrt{4 - x^2}\right]^2\right] - \left[\left[\sqrt{4 - x^2}\right]^2 - \left[-\sqrt{4 - x^2}\right]^2\right]\right] dx$$

$$\int_C (x^2 + y^2)dx + 3x^2ydy = \int_{x=-2}^{2} [3x[(4 - x^2) - (4 - x^2)] - [(4 - x^2) - (4 - x^2)]]dx$$

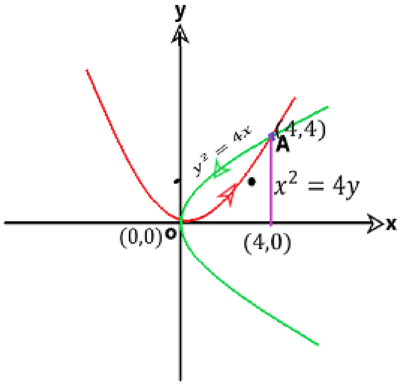
$\boldsymbol{\int_C (x^2 + y^2)dx + 3x^2ydy = 0}$.

3.Use Green's theorem to find area between the parabola $x^2 = 4y$ and $y^2 = 4x$.

Sol: Wkt **Area** $= \iint_R \boldsymbol{dydx} = \frac{1}{2}\int_C (\boldsymbol{xdy - ydx})$

Thus, $\quad A = \frac{1}{2}\int_C (xdy - ydx)$

$A = \frac{1}{2}\left[\int_{OA}(xdy - ydx) + \int_{AO}(xdy - ydx)\right]$ --- (1)



$\boldsymbol{y^2 = 4x}$ & $\boldsymbol{x^2 = 4y}$

$\boldsymbol{y^4 = 16x^2}$

$\boldsymbol{y^4 = 16(4y)}$

$\boldsymbol{y^4 = 64y}$

$\boldsymbol{y^4 - 64y = 0}$

$\boldsymbol{y(y^3 - 64) = 0}$

$\boldsymbol{y = 0 \quad y^3 = 64}$

$\boldsymbol{y = 0}$ & $\boldsymbol{y = 4}$

Along OA:	Along AO:
$\boldsymbol{x^2 = 4y}$	$\boldsymbol{y^2 = 4x}$
$y = \frac{x^2}{4}$	$x = \frac{y^2}{4}$
$dy = \frac{2x}{4}dx$	$dx = \frac{2y}{4}dy$
$\boldsymbol{dy = \frac{x}{2}dx}$	$\boldsymbol{dx = \frac{y}{2}dy}$

In 'R'

$$x: x = 0 \ \ to \ \ x = 4y: y = 4 \ \ to \ \ y = 0$$

$$(1) \Rightarrow A = \frac{1}{2}\left[\int_{x=0}^{4}\left(x.\frac{x}{2}dx - \frac{x^2}{4}dx\right) + \int_{y=4}^{0}\left(\frac{y^2}{4}dy - y.\frac{y}{2}dy\right)\right]$$

$$A = \frac{1}{2}\left[\int_{x=0}^{4}\left(\frac{x^2}{2} - \frac{x^2}{4}\right)dx + \int_{y=4}^{0}\left(\frac{y^2}{4} - \frac{y^2}{2}\right)dy\right]$$

$$A = \frac{1}{2}\left[\int_{x=0}^{4}\left(\frac{x^2}{4}\right)dx - \int_{y=4}^{0}\left(\frac{y^2}{4}\right)dy\right]$$

$$A = \frac{1}{2*4}\left[\int_{x=0}^{4} x^2\,dx - \int_{y=4}^{0} y^2\,dy\right]$$

$$A = \frac{1}{8}\left[\left[\frac{x^3}{3}\right]_0^4 - \left[\frac{y^3}{3}\right]_4^0\right]$$

$$A = \frac{1}{8*3}[(4^3 - 0) - (0 - 4^3)]$$

$$A = \frac{1}{24}[64 + 64]$$

$$\mathrm{A} = \frac{128}{24}$$

$\boldsymbol{A} = \frac{\mathbf{16}}{\mathbf{3}}$ **Sq.Units**

Stoke's Theorem

If S is a surface bounded by a simple closed curve C and if $\vec{F}$ is any continuously differentiable vector function then

$$\oint_C \vec{\boldsymbol{F}}.\boldsymbol{d}\vec{\boldsymbol{r}} = \iint_S \boldsymbol{curl}\vec{\boldsymbol{F}}.\hat{\boldsymbol{n}}\,\boldsymbol{ds} = \iint_S (\boldsymbol{\nabla X}\vec{\boldsymbol{F}}).\hat{\boldsymbol{n}}\,\boldsymbol{ds}$$

Problems

1.Verify Stoke's theorem for $\vec{F} = y\hat{\imath} + z\hat{\jmath} + x\hat{k}$ where S is the upper half of the sphere

$x^2 + y^2 + z^2 = 1$ and C is its boundary.

Sol: By Stoke's theorem,

$\oint_C \vec{\boldsymbol{F}}.\boldsymbol{d}\vec{\boldsymbol{r}} = \iint_S \boldsymbol{curl}\vec{\boldsymbol{F}}.\hat{\boldsymbol{n}}\,\boldsymbol{ds}$, C is the circle in the $xy - plane$ whose centre is the origin and radius equal to unity.

i.e., $x^2 + y^2 = 1$ and $z = 0$

$dz = 0$

Put $x = cos\theta \; y = sin\theta$; $0 \le \theta \le 2\pi$

LHS, $\oint_C \vec{F}.d\vec{r} = \oint_C ydx + zdy + xdz$

$= \int_{\theta=0}^{2\pi} (sin\theta)(-sin\theta d\theta) + (0) + (0)$

$= -\int_{\theta=0}^{2\pi} sin^2\theta d\theta$

$= -\int_{\theta=0}^{2\pi} \frac{(1-cos2\theta)}{2} d\theta$

$= -\frac{1}{2}\int_{\theta=0}^{2\pi} (1 - cos2\theta)d\theta$

$= -\frac{1}{2}\left\{[\theta]_0^{2\pi} - \left[\frac{sin2\theta}{2}\right]_0^{2\pi}\right\}$

$= -\frac{1}{2}\left\{[2\pi - 0] - \frac{1}{2}[sin4\pi - sin0]\right\}$

$$\oint_C \vec{F}.d\vec{r} = -\frac{1}{2}(2\pi)$$

$$\oint_C \vec{\boldsymbol{F}}.\boldsymbol{d}\vec{\boldsymbol{r}} = -\boldsymbol{\pi}$$

Now, $curl\vec{F} = \nabla X\vec{F}$

$$= \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y & z & x \end{vmatrix}$$

$$= \left[\frac{\partial}{\partial y}(x) - \frac{\partial}{\partial z}(z)\right]\hat{\imath} - \left[\frac{\partial}{\partial x}(x) - \frac{\partial}{\partial z}(y)\right]\hat{\jmath} + \left[\frac{\partial}{\partial x}(z) - \frac{\partial}{\partial y}(y)\right]\hat{k}$$

$$= [0-1]\hat{\imath} - [1-0]\hat{\jmath} + [0-1]\hat{k}$$

$$curl\vec{F} = -\hat{\imath} - \hat{\jmath} - \hat{k}$$

$$\hat{n}ds = dydz\hat{\imath} + dxdz\hat{\jmath} + dxdy\hat{k}$$

$\hat{n}ds = 0.\hat{\imath} + 0.\hat{\jmath} + dxdy\hat{k}$ $\qquad$ $(Since\ z = 0,\ dz = 0)$

$$curl\vec{F}.\hat{n}ds = \left(-\hat{\imath} - \hat{\jmath} - \hat{k}\right).\ (0.\hat{\imath} + 0.\hat{\jmath} + dxdy\hat{k})$$

$$\boldsymbol{curl\vec{F}.\hat{n}ds = -dxdy}$$

RHS, $\iint_S \ curl\vec{F}\,.\hat{n}\ ds = \iint_S \ -dxdy$

$$\iint_S curl\vec{F}\,.\hat{n}\ ds = -\iint_S dxdy$$

$\iint_S \ \boldsymbol{curl\vec{F}\,.\hat{n}\ ds = -\pi}$ $\qquad$ $\left(\boldsymbol{Since} \iint_S \ dxdy = \boldsymbol{Area\ of\ circle}, x^2 + y^2 = 1 = \pi(1)^2 = \pi\right)$

LHS = RHS

2. Evaluate $\oint_C \ (xydx + xy^2dy)$ by Stoke's theorem where C is the square in $xy-\ plane$ with vertices $(1,0)\ (-1,0)\ \ (0,1)\ (0,-1)$.

Sol: Given $\oint_C \ (xydx + xy^2dy)$

$$\oint_C \vec{F}.d\vec{r} = \oint_C (xydx + xy^2dy)$$

Here, $\vec{F} = xy\hat{\imath} + xy^2\hat{\jmath} + 0\hat{k}$

Now, $curl\vec{F} = \nabla X\vec{F}$

$$= \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ xy & xy^2 & 0 \end{vmatrix}$$

$$= \left[\frac{\partial}{\partial y}(0) - \frac{\partial}{\partial z}(xy^2)\right]\hat{\imath} - \left[\frac{\partial}{\partial x}(0) - \frac{\partial}{\partial z}(xy)\right]\hat{\jmath} + \left[\frac{\partial}{\partial x}(xy^2) - \frac{\partial}{\partial y}(xy)\right]\hat{k}$$

$$curl\vec{F} = [0-0]\hat{\imath} - [0-0]\hat{\jmath} + [y^2 - x]\hat{k}$$

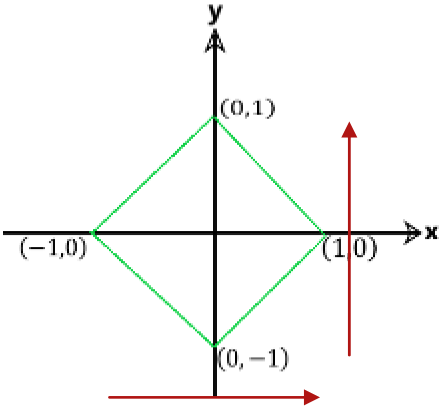
$$\hat{n}ds = dydz\hat{\imath} + dxdz\hat{\jmath} + dxdy\hat{k}$$

$$curl\vec{F}.\hat{n}ds = \left(0.\hat{\imath} - 0.\hat{\jmath} + [y^2 - x]\hat{k}\right).\ (dydz\hat{\imath} + dxdz\hat{\jmath} + dxdy\hat{k})$$

$$\boldsymbol{curl\vec{F}.\hat{n}ds = [y^2 - x]dxdy}$$

Wkt , $\oint_C \boldsymbol{\vec{F}.d\vec{r}} = \iint_R \boldsymbol{curl\vec{F}\,.\hat{n}\,ds}$

$$= \iint_R [y^2 - x]dxdy$$



In ‘R’

$x: x = -1 \;\; to \;\; x = 1$

$y: y = -1 \;\; to \;\; y = 1$

$$\oint_C \boldsymbol{\vec{F}.d\vec{r}} = \int_{x=-1}^{1}\int_{y=-1}^{1} [y^2 - x]dydx$$

$$= \int_{x=-1}^{1} \left\{\left[\frac{y^3}{3}\right]_{-1}^{1} - x[y]_{-1}^{1}\right\} dx$$

$$= \int_{x=-1}^{1} \left[\frac{1}{3}[1-(-1)] - x[1-(-1)]\right] dx$$

$$\oint_C \vec{F}.d\vec{r} = \int_{x=-1}^{1} \left[\frac{2}{3} - 2x\right] dx$$

$$\oint_C \vec{F}.d\vec{r} = \frac{2}{3}[x]_{-1}^{1} - 2\left[\frac{x^2}{2}\right]_{-1}^{1}$$

$$\oint_C \vec{F}.d\vec{r} = \frac{2}{3}[1-(-1)] - [(1)^2 - (-1)^2]$$

$$\oint_C \vec{F}.d\vec{r} = \frac{2}{3}[2] - [0]$$

$$\oint_C \boldsymbol{\vec{F}.d\vec{r}} = \boldsymbol{\frac{4}{3}}$$

3. Evaluate $\oint_C (x^2 + y^2)dx - 2xydy$ taken round the rectangle bounded by $x = 0, x = a,$ $y = 0, y = b$ using Stoke’s theorem .

Sol: Given $\oint_C (x^2 + y^2)dx - 2xydy$

$$\oint_C \vec{F}.d\vec{r} = \oint_C (x^2+y^2)dx - 2xydy$$

Here, $\vec{F} = (x^2+y^2)\hat{\imath} - 2xy\hat{\jmath} + 0\hat{k}$

Now, $curl\vec{F} = \nabla X\vec{F}$

$$= \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ (x^2+y^2) & -2xy & 0 \end{vmatrix}$$

$$= \left[\frac{\partial}{\partial y}(0) - \frac{\partial}{\partial z}(-2xy)\right]\hat{\imath} - \left[\frac{\partial}{\partial x}(0) - \frac{\partial}{\partial z}((x^2+y^2))\right]\hat{\jmath} + \left[\frac{\partial}{\partial x}(-2xy) - \frac{\partial}{\partial y}((x^2+y^2))\right]\hat{k}$$

$$curl\vec{F} = [0-0]\hat{\imath} - [0-0]\hat{\jmath} + [-2y-2y]\hat{k}$$

$$curl\vec{F} = 0\hat{\imath} - 0\hat{\jmath} - 4y\hat{k}$$

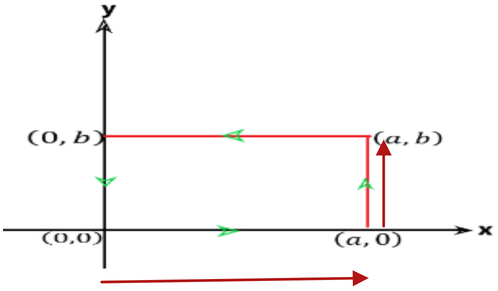
$$\hat{n}ds = dydz\hat{\imath} + dxdz\hat{\jmath} + dxdy\hat{k}$$

$$curl\vec{F}.\hat{n}ds = \left(0.\hat{\imath} - 0.\hat{\jmath} - 4y\hat{k}\right).\ (dydz\hat{\imath} + dxdz\hat{\jmath} + dxdy\hat{k})$$

$$\boldsymbol{curl\vec{F}.\hat{n}ds = -4ydxdy}$$

Wkt , $\oint_C \boldsymbol{\vec{F}.d\vec{r}} = \iint_R \boldsymbol{curl\vec{F}.\hat{n}\,ds}$

$$= \iint_R -4ydxdy$$



In 'R'

$x: x = 0\ \ to\ \ x = a$

$y: y = 0\ \ to\ \ y = b$

$$\oint_C \boldsymbol{\vec{F}.d\vec{r}} = -4\int_{x=0}^{a}\int_{y=0}^{b} y\,dy\,dx$$

$$= -4\int_{x=0}^{a}\left\{\left[\frac{y^2}{2}\right]_0^b\right\}dx$$

$$= -4\int_{x=0}^{a}\left\{\left[\frac{b^2}{2} - 0\right]\right\}dx$$

$$\oint_C \vec{F}.d\vec{r} = -4.\frac{b^2}{2}\int_{x=0}^{a} 1.dx$$

$$\oint_C \vec{F}.d\vec{r} = -2b^2[x]_0^a$$

$$\oint_C \vec{F}.d\vec{r} = -2b^2(a-0)$$

$$\boldsymbol{\oint_C \vec{F}.d\vec{r} = -2ab^2}$$

4. Using Stoke's theorem to evaluate $\oint_C \vec{F}.d\vec{r}$ where $\vec{F} = y\hat{\imath} + z\hat{\jmath} + x\hat{k}$ and C is the boundary of the upper half of the sphere $x^2 + y^2 + z^2 = 1$.

Sol: Given $\vec{F} = y\hat{\imath} + z\hat{\jmath} + x\hat{k}$

C is the circle in the $xy - plane$ whose centre is the origin and radius equal to unity.

i.e., $x^2 + y^2 = 1$ and $z = 0$

$dz = 0$

Now, $curl\vec{F} = \nabla X\vec{F}$

$$= \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y & z & x \end{vmatrix}$$

$$= \left[\frac{\partial}{\partial y}(x) - \frac{\partial}{\partial z}(z)\right]\hat{\imath} - \left[\frac{\partial}{\partial x}(x) - \frac{\partial}{\partial z}(y)\right]\hat{\jmath} + \left[\frac{\partial}{\partial x}(z) - \frac{\partial}{\partial y}(y)\right]\hat{k}$$

$$= [0-1]\hat{\imath} - [1-0]\hat{\jmath} + [0-1]\hat{k}$$

$$curl\vec{F} = -\hat{\imath} - \hat{\jmath} - \hat{k}$$

$$\hat{n}ds = dydz\hat{\imath} + dxdz\hat{\jmath} + dxdy\hat{k}$$

$\hat{n}ds = 0.\hat{\imath} + 0.\hat{\jmath} + dxdy\hat{k}$ $\qquad (Since\ z = 0,\ dz = 0)$

$$curl\vec{F}.\hat{n}ds = \left(-\hat{\imath} - \hat{\jmath} - \hat{k}\right).\ (0.\hat{\imath} + 0.\hat{\jmath} + dxdy\hat{k})$$

$$\boldsymbol{curl\vec{F}.\hat{n}ds = -dxdy}$$

By Stoke's theorem,

$$\boldsymbol{\oint_C \vec{F}.d\vec{r} = \iint_S curl\vec{F}.\hat{n}\,ds}$$

$$\oint_C \vec{F}.d\vec{r} = \iint_S -dxdy$$

$$\oint_C \vec{F}.d\vec{r} = -\iint_S dxdy$$

$\oint_C \boldsymbol{\vec{F}.d\vec{r}} = -\boldsymbol{\pi}$ $\left(Since \iint_S \ dxdy = Area\ of\ circle, x^2+y^2=1=\pi(1)^2=\pi\right)$

Problem on flux

If $\vec{F} = 2xy\hat{\imath} + yz^2\hat{\jmath} + xz\hat{k}$ and S is the rectangular parallelepiped bounded by $x=0, y=0, z=0,\ x=2, y=1, z=3$.Find the flux across S.

Sol: Here$\vec{F} = 2xy\hat{\imath} + yz^2\hat{\jmath} + xz\hat{k}$

Now, $div\vec{F} = \nabla.\vec{F}$

$$= \left[\frac{\partial}{\partial x}\hat{\imath} + \frac{\partial}{\partial y}\hat{\jmath} + \frac{\partial}{\partial z}\hat{k}\right].\left[2xy\hat{\imath} + yz^2\hat{\jmath} + xz\hat{k}\right]$$

$$div\vec{F} = \frac{\partial}{\partial x}(2xy) + \frac{\partial}{\partial y}(yz^2) + \frac{\partial}{\partial z}(xz)$$

$$\boldsymbol{div\vec{F} = 2y + z^2 + x}$$

Flux across S $= \iint_S \ \vec{F}.\hat{n}\,ds$

By divergence theorem, $\iint_S \ \vec{F}.\hat{n}\,ds = \iiint_V \ div\vec{F}\,dv$

$$\iint_S \vec{F}.\hat{n}\,ds = \int_{z=0}^{3}\int_{y=0}^{1}\int_{x=0}^{2}(\boldsymbol{2y + z^2 + x})\,dx\,dy\,dz$$

$$= \int_{z=0}^{3}\int_{y=0}^{1}\left[(2y+z^2)\,[x]_0^2 + \left[\frac{y^2}{2}\right]_0^2\right]dy\,dz$$

$$= \int_{z=0}^{3}\int_{y=0}^{1}(4y + 2z^2 + 2)\,dy\,dz$$

$$= \int_{z=0}^{3}\left\{4\left[\frac{y^2}{2}\right]_0^1 + (2z^2+2)[y]_0^1\right\}dz$$

$$= \int_{z=0}^{3}\{2 + 2z^2 + 2\}dz$$

$$= \int_{z=0}^{3}\{2z^2 + 4\}dz$$

$$\iint_S \vec{F}.\hat{n}\,ds = 2\left[\frac{z^3}{3}\right]_0^3 + 4[z]_0^3$$

$$= 2(9-0) + 4(3-0)$$

$$\iint_{\boldsymbol{S}} \boldsymbol{\vec{F}.\hat{n}\,ds = 30}$$

1 VECTOR SPACES AND SUBSPACES

What is a vector? Many are familiar with the concept of a vector as:

- Something which has magnitude and direction.
- an ordered pair or triple.
- a description for quantities such as Force, velocity and acceleration.

Such vectors belong to the foundation vector space - $\mathbb{R}^n$ - of all vector spaces. The properties of general vector spaces are based on the properties of $\mathbb{R}^n$. It is therefore helpful to consider briefly the nature of $\mathbb{R}^n$.

1.1 The Vector Space $\mathbb{R}^n$

Definitions

- If n is a positive integer, then an **ordered n-tuple** is a sequence of n real numbers $(a_1, a_2, \ldots, a_n)$. The set of all ordered n-tuples is called **n-space** and is denoted by $\mathbb{R}^n$.

When $n = 1$ each ordered n-tuple consists of one real number, and so $\mathbb{R}$ may be viewed as the set of real numbers. Take $n = 2$ and one has the set of all 2-tuples which are more commonly known as **ordered pairs**. This set has the geometrical interpretation of describing all points and directed line segments in the Cartesian $x-y$ plane. The vector space $\mathbb{R}^3$, likewise is the set of **ordered triples**, which describe all points and directed line segments in 3-D space.

In the study of 3-space, the symbol (a_1, a_2, a_3) has two different geometric interpretations: it can be interpreted as a point, in which case a_1, a_2 and a_3 are the coordinates, or it can be interpreted as a vector, in which case a_1, a_2 and a_3 are the components. It follows, therefore, that an ordered n-tuple $(a_1, a_2, \ldots, a_n)$ can be viewed as a "generalized point" or a "generalized vector" - the distinction is mathematically unimportant. Thus, we can describe the 5-tuple $(1,2,3,4,5)$ either as a point or a vector in $\mathbb{R}^5$.

Definitions

- Two vectors $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ in $\mathbb{R}^n$ are called **equal** if
$$u_1 = v_1, u_2 = v_2, \ldots, u_n = v_n$$
- The **sum** $u + v$ is defined by
$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$$
- Let k be any scalar, then the **scalar multiple** $k\mathbf{u}$ is defined by
$$k\mathbf{u} = (ku_1, ku_2, \ldots, ku_n)$$
- These two operations of addition and scalar multiplication are called the **standard operations** on $\mathbb{R}^n$.
- The **zero vector** in $\mathbb{R}^n$ is denoted by $\mathbf{0}$ and is defined to be the vector
$$\mathbf{0} = (0, 0, \ldots, 0)$$
- The **negative** (or **additive inverse**) of u is denoted by -u and is defined by
$$-\mathbf{u} = (-u_1, -u_2, \ldots, -u_n)$$
- The **difference** of vectors in $\mathbb{R}^n$ is defined by
$$\mathbf{v} - \mathbf{u} = \mathbf{v} + (-\mathbf{u})$$

The most important arithmetic properties of addition and scalar multiplication of vectors in $\mathbb{R}^n$ are listed in the following theorem. This theorem enabes us to manipulate vectors in $\mathbb{R}^n$ without expressing the vectors in terms of componenets.

Theorem 1.1. *If* $\mathbf{u} = (u_1, u_2, \ldots, u_n), \mathbf{v} = (v_1, v_2, \ldots, v_n)$*, and* $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ *are vectors in* $\mathbb{R}^n$ *and* k *and* l *are scalars, then:*

1. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
2. $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
3. $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$
4. $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$*; that is,* $\mathbf{u} - \mathbf{u} = \mathbf{0}$
5. $k(l\mathbf{u}) = (kl)\mathbf{u}$
6. $k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$
7. $(k + l)\mathbf{u} = k\mathbf{u} + l\mathbf{u}$
8. $1\mathbf{u} = \mathbf{u}$

1.2 Generalized Vector Spaces

The time has now come to generalize the concept of a vector. In this section a set of axioms are stated, which if satisfied by a class of objects, entitles those objects to be called “vectors”. The axioms were chosen by abstracting the most important properties (theorem 1.1). of vectors in $\mathbb{R}^n$; as a consequence, vectors in $\mathbb{R}^n$ automatically satisfy these axioms. Thus, the new concept of a vector, includes many new kinds of vector without excluding the “common vector”. The new types of vectors include, among other things, various kinds of matrices and functions.

Definition

A *vector space* V over a field $\mathbb{F}$ is a nonempty set on which two operations are defined - addition and scalar multiplication. Addition is a rule for associating with each pair of objects $\mathbf{u}$ and $\mathbf{v}$ in V an object $\mathbf{u} + \mathbf{v}$, and scalar multiplication is a rule for associating with each scalar $k \in \mathbb{F}$ and each object $\mathbf{u}$ in V an object $k\mathbf{u}$ such that

1. If $\mathbf{u}, \mathbf{v} \in V$, then $\mathbf{u} + \mathbf{v} \in V$.
2. If $\mathbf{u} \in V$ and $k \in \mathbb{F}$, then $k\mathbf{u} \in V$.
3. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
4. $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
5. There is an object $\mathbf{0}$ in V, called a **zero vector** for V, such that $\mathbf{u}+\mathbf{0} = \mathbf{0}+\mathbf{u} = \mathbf{u}$ for all $\mathbf{u}$ in V.
6. For each $\mathbf{u}$ in V, there is an object **-u** in V, called the **additive inverse** of $\mathbf{u}$, such that $\mathbf{u} + (-\mathbf{u}) = -\mathbf{u} + \mathbf{u} = \mathbf{0}$;
7. $k(l\mathbf{u}) = (kl)\mathbf{u}$
8. $k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$
9. $(k + l)\mathbf{u} = k\mathbf{u} + l\mathbf{u}$
10. $1\mathbf{u} = \mathbf{u}$

Remark The elements of the underlying field $\mathbb{F}$ are called scalars and the elements of the vector space are called vectors. Note also that we often restrict our attention to the case when $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

Examples of Vector Spaces

A wide variety of vector spaces are possible under the above definition as illustrated by the following examples. In each example we specify a nonempty set of objects V. We must then define two operations - addition and scalar multiplication, and as an exercise we will demonstrate that all the axioms are satisfied, hence entitling V with the specified operations, to be called a vector space.

1. The set of all n-tuples with entries in the field $\mathbb{F}$, denoted $\mathbb{F}^n$ (especially note $\mathbb{R}^n$ and $\mathbb{C}^n$).

2. The set of all $m \times n$ matrices with entries from the field $\mathbb{F}$, denoted $M_{m\times n}(\mathbb{F})$.

3. The set of all real-valued functions defined on the real line $(-\infty, \infty)$.

4. The set of polynomials with coefficients from the field $\mathbb{F}$, denoted $P(\mathbb{F})$.

5. (Counter example) Let $V = \mathbb{R}^2$ and define addition and scalar multiplication oparations as follows: If $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$, then define

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2)$$

and if k is any real number, then define

$$k\mathbf{u} = (ku_1, 0).$$

1.2.1 Some Properties of Vectors

It is important to realise that the following results hold for all vector spaces. They provide a useful set of vector properties.

Theorem 1.2. *If $u, v, w \in V$ (a vector space) such that $u + w = v + w$, then $u = v$.*

Corollary 1.1. *The zero vector and the additive inverse vector (for each vector) are unique.*

Theorem 1.3. *Let V be a vector space over the field $\mathbb{F}$, $\mathbf{u} \in V$, and $k \in \mathbb{F}$. Then the following statement are true:*

(a) $0\mathbf{u} = \mathbf{0}$

(b) $k\mathbf{0} = \mathbf{0}$

(c) $(-k)\mathbf{u} = -(k\mathbf{u}) = k(-\mathbf{u})$

(d) If $k\mathbf{u} = \mathbf{0}$*, then* $k = 0$ *or* $\mathbf{u} = 0$.

1.2.2 Quiz

True or false?

(a) Every vector space contains a zero vector.

(b) A vector space may have more than one zero vector.

(c) In any vector space, $a\mathbf{u} = b\mathbf{u}$ implies $a = b$.

(d) In any vector space, $a\mathbf{u} = a\mathbf{v}$ implies $\mathbf{u} = \mathbf{v}$.

1.3 Subspaces

It is possible for one vector space to be contained within a larger vector space. This section will look closely at this important concept.

Definitions

- A subset W of a vector space V is called a **subspace** of V if W is itself a vector space under the addition and scalar multiplication defined on V.

In general, all ten vector space axioms must be verified to show that a set W with addition and scalar multiplication forms a vector space. However, if W is part of a larget set V that is already known to be a vector space, then certain axioms need not be verified for W because they are inherited from V. For example, there is no need to check that $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (axiom 3) for W because this holds for all vectors in V and consequently holds for all vectors in W. Likewise, axioms 4, 7, 8, 9 and 10 are inherited by W from V. Thus to show that W is a subspace of a vector space V (and hence that W is a vector space), only axioms 1, 2, 5 and 6 need to be verified. The following theorem reduces this list even further by showing that even axioms 5 and 6 can be dispensed with.

Theorem 1.4. *If W is a set of one or more vectors from a vector space V, then W is a subspace of V if and only if the following conditions hold.*

(a) If $\mathbf{u}$ and $\mathbf{v}$ are vectors in W, then $\mathbf{u} + \mathbf{v}$ is in W.

(b) If k is any scalar and $\mathbf{u}$ is any vector in W, then $k\mathbf{u}$ is in W.

Proof. If W is a subspace of V, then all the vector space axioms are satisfied; in particular, axioms 1 and 2 hold. These are precisely conditions (a) and (b).

Conversely, assume conditions (a) and (b) hold. Since these conditions are vector space axioms 1 and 2, it only remains to be shown that W satisfies the remaining eight axioms. Axioms 3, 4, 7, 8, 9 and 10 are automatically satisfied by the vectors in W since they are satisfied by all vectors in V. Therefore, to complete the proof, we need only verify that Axioms 5 and 6 are satisfied by vectors in W.

Let $\mathbf{u}$ be any vector in W. By condition (b), k$\mathbf{u}$ is in W for every scalar k. Setting $k = 0$, it follows from theorem 1.3 that $0\mathbf{u} = \mathbf{0}$ is in W, and setting $k = -1$, it follows that $(-1)\mathbf{u} = -\mathbf{u}$ is in W. □

Remarks

- Note that a consequence of (b) is that $\mathbf{0}$ is an element of W.
- A set W of one or more vectors from a vector space V is said to be **closed under addition** if condition (a) in theorem 1.4 holds and **closed under scalar multiplication** if condition (b) holds. Thus, theorem 1.4 states that W is *a* subspace of V if and only if W is closed under addition and closed under scalar multiplication.

Examples of Subspaces

1. A plane through the origin of $\mathbb{R}^3$ forms a subspace of $\mathbb{R}^3$. This is evident geometrically as follows: Let W be any plane through the origin and let $\mathbf{u}$ and $\mathbf{v}$ be any vectors in W other than the zero vector. Then $\mathbf{u} + \mathbf{v}$ must lie in W because it is the diagonal of the parallelogram determined by $\mathbf{u}$ and $\mathbf{v}$, and $k\mathbf{u}$ must lie in W for any scalar k because $k\mathbf{u}$ lies on a line through $\mathbf{u}$. Thus, W is closed under addition and scalar multiplication, so it is a subspace of $\mathbb{R}^3$.

2. A line through the origin of $\mathbb{R}^3$ is also a subspace of $\mathbb{R}^3$. It is evident geometrically that the sum of two vectors on this line also lies on the line and that a scalar multiple of a vector on the line is on the line as well. Thus, W is closed under addition and scalar multiplication, so it is a subspace of $\mathbb{R}^3$.

3. Let n be a positive integer, and let W consist of all functions expressible in the form

   $$p(x) = a_0 + a_1x + \cdots + a_nx^n$$

   where $a_0, \ldots, a_n$ belong to some field $\mathbb{F}$. Thus, W consists of the zero function together with all polynomials in $\mathbb{F}$ of degree n or less. The set W is a subspace of $P(\mathbb{F})$ (example 4 on page 5), and if $\mathbb{F} = \mathbb{R}$ it is also a subspace of the vector space of all real-valued functions (discussed in example 3 on page 5).

   To see this, let $\mathbf{p}$ and $\mathbf{q}$ be the polynomials

   $$p(x) = a_0 + a_1x + \cdots + a_nx^n$$

   and

   $$q(x) = b_0 + b_1x + \cdots + b_nx^n$$

   Then

   $$(\mathbf{p} + \mathbf{q})(x) = p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n$$

   and

   $$(k\mathbf{p})(x) = kp(x) = (ka_0) + (ka_1)x + \cdots + (ka_n)x^n$$

   These functions have the form given above, so $\mathbf{p} + \mathbf{q}$ and $k\mathbf{p}$ lie in W. This vector space W is denoted $P_n(\mathbb{F})$.

4. The *transpose* A^T of an $m \times n$ matrix A is the $n \times m$ matrix obtained from A by interchanging rows and columns. A *symmetric matrix* is a square matrix A such that $A^T = A$. The set of all symmetric matrices in $M_{n\times n}(\mathbb{F})$ is a subspace of $M_{n\times n}(\mathbb{F})$.

5. The trace of an $n \times n$ matrix A, denoted $\mathrm{tr}(A)$, is the sum of the diagonal entries of A. The set of $n \times n$ matrices having trace equal to zero is a subspace of $M_{n\times n}(\mathbb{F})$.

1.3.1 Operations on Vector Spaces

Definitions

- The addition of two subsets U and V of a vector space is defined by:

$$U + V = \{\mathbf{u} + \mathbf{v} | \mathbf{u} \in U, \mathbf{v} \in V\}$$

- The intersection $\cap$ of two subsets U and V of a vector space is defined by:

$$U \cap V = \{\mathbf{w} | \mathbf{w} \in U \; and \; \mathbf{w} \in V\}$$

- A vector space W is called the direct sum of U and V, denoted $U \oplus V$, if U and V are subspaces of W with $U \cap V = \{0\}$ and $U + V = W$.

The following theorem shows how we can form a new subspace from other ones.

Theorem 1.5. *Any intersection or sum of subspaces of a vector space V is also a subspace of V.*

1.3.2 Quiz

True or false?

(a) If V is a vector space and W is a subset of V that is also a vector space, then W is a subspace of V.

(b) The empty set is a subspace of every vector space.

(c) If V is a vector space other than the zero vector space, then V contains a subspace W such that $W \neq V$.

(d) The intersection of any two subsets of V is a subspace of V.

(e) Any union of subspaces of a vector space V is a subspace of V.

1.4 Linear Combinations of Vectors and Systems of Linear Equations

Have m linear equations in n variables:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

Write in matrix form: $A\boldsymbol{x} = \boldsymbol{b}$.

$A = [a_{ij}]$ is the $m \times n$ coefficient matrix.

$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ is the column vector of unknowns, and $\boldsymbol{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ is the column vector of RHS.

Note: a_{ij}, $b_j \in \mathbb{R}$ or $\mathbb{C}$.

1.4.1 Gaussian Elimination

To solve $A\boldsymbol{x} = \boldsymbol{b}$:

write *augmented matrix*: $[A|\boldsymbol{b}]$.

1. Find the left-most non-zero column, say column j.

2. Interchange top row with another row if necessary, so top element of column j is non-zero. (The **pivot**.)

3. Subtract multiples of row 1 from all other rows so all entries in column j below the top are then 0.

4. Cover top row; repeat **1** above on rest of rows.

Continue until all rows are covered, or until only $00\ldots0$ rows remain.

Result is a triangular system, easily solved by *back substitution*: solve the last equation first, then 2nd last equation and so on.

1.4.2 Example

Use Gaussian elimination to solve:

$$\begin{aligned} x_3 - \quad x_4 &= 2 \\ -9x_1 - 2x_2 + 6x_3 - 12x_4 &= -7 \\ 3x_1 + \ x_2 - 2x_3 + \ 4x_4 &= 2 \\ 2x_3 \qquad\quad &= 6 \end{aligned}$$

1.4.3 Definition (row echelon form)

A matrix is in *row echelon form* (r.e.f.) if each row after the first starts with *more* zeros than the previous row (or else rows at bottom of matrix are all zeros).

The Gauss algorithm converts any matrix to one in row echelon form. The 2 matrices are *equivalent*, that is, they have the same solution set.

1.4.4 Elementary row operations

1. $r_i \leftrightarrow r_j$: swap rows i and j.

2. $r_i \rightarrow r_i - cr_j$: replace row i with
(row i minus c times row j).

3. $r_i \rightarrow cr_i$:
replace row i with c times row i, where $c \neq 0$.

The Gauss algorithm uses only **1** and **2**.

1.4.5 Possible solutions for $A\boldsymbol{x} = \boldsymbol{b}$

Consider the r.e.f. of $[A|\boldsymbol{b}]$. Then we have three possibilities:

(1) *Exactly one* solution; here the r.e.f. gives each variable a single value, so the number of variables, n, equals the number of non-zero rows in the r.e.f.

(2) *No* solution; when one row of r.e.f. is $(0\,0\,\ldots\,d)$ with $d \neq 0$. We can't solve $0x_1 + 0x_2 + \cdots + 0x_m = d$ if $d \neq 0$; it says $0 = d$. In this case the system is said to be *inconsistent*.

(3) *Infinitely many* solutions; here the number of rows of the r.e.f. is *less* than the number of variables.

Note that a *homogeneous* system has $\boldsymbol{b} = \mathbf{0}$, i.e., all zero RHS. Then we always have at least the trivial solution, $x_i = 0$, $1 \leq i \leq n$.

1.4.6 Examples

$$\begin{aligned} x_1 + x_2 - x_3 &= 0 \\ 2x_1 - x_2 \phantom{{}-2x_3} &= 0 \\ 4x_1 + x_2 - 2x_3 &= 0 \end{aligned}$$

$$\begin{aligned} x_2 - 2x_3 + 4x_4 &= 2 \\ 2x_2 - 3x_3 + 7x_4 &= 6 \\ x_3 - x_4 &= 2 \end{aligned}$$

1.4.7 Different right hand sides

To solve $A\boldsymbol{x} = \boldsymbol{b}_j$, for $j = 1, \ldots, r$, for r different sets of right hand sides $\boldsymbol{b}_j$:

Form a *big* augmented matrix $[A|\boldsymbol{b}_1\boldsymbol{b}_2\ldots\boldsymbol{b}_r]$ and find its r.e.f. $[U|\boldsymbol{b}'_1\boldsymbol{b}'_2\ldots\boldsymbol{b}'_r]$. So U will be a r.e.f. corresponding to A. Then solve each of the systems $U\boldsymbol{x} = \boldsymbol{b}'_j$, $j = 1, 2, \ldots, r$, by back substitution.

1.4.8 Special case: finding A^{-1} (if it exists)

If A is $n \times n$ and it has an inverse, then solving $A\boldsymbol{x} = \boldsymbol{e}_j$ (where $\boldsymbol{e}_j$ is the $n \times 1$ column with 1 in jth place and 0 elsewhere) gives jth column of A^{-1}.

So we find r.e.f. of $[A|\boldsymbol{e}_1\boldsymbol{e}_2 \dots \boldsymbol{e}_n]$, i.e., determine the r.e.f. of $[A|I]$ where I is $n \times n$ identity matrix.

Once we have found the r.e.f. of $[A|I]$ to be $[U|*]$, we then use row operations to convert it to $[I|D]$, so $D = A^{-1}$.

If the last row of U is all zeros, A has no inverse.

Note that if A and I are square, $AC = I$ implies $CA = I$ and conversely.

If such a matrix C exists, it is unique. We write $C = A^{-1}$, and we say A is *non-singular* or *invertible*.

1.4.9 Example

Does $A = \begin{pmatrix} 1 & -1 & 4 \\ 1 & 0 & -2 \\ 2 & -2 & 10 \end{pmatrix}$ have an inverse?

If so, find it.

1.4.10 Linear combinations

Definitions

- A vector $\mathbf{w}$ is called a **linear combination** of the vectors $v_1, v_2, \dots, v_r$ if it can be expressed in the form

$$\mathbf{w} = k_1 v_1 + k_2 v_2 + \cdots + k_r v_r$$

where $k_1, k_2, \dots, k_r$ are scalars.

Example

1. Consider the vectors $\mathbf{u} = (1, 2, -1)$ and $\mathbf{v} = (6, 4, 2)$ in $\mathbb{R}^3$. Show that $\mathbf{w} = (9, 2, 7)$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$ and that $\mathbf{w'} = (4, -1, 8)$ is not a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

1.4.11 Spanning

If $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_r}$ are vectors in a vector space V, then generally some vectors in V may be linear combinations of $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_r}$ and others may not. The following theorem shows that if a set W is constructed consisting of all those vectors that are expressible as linear combinations of $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_r}$, then W forms a subspace of V.

Theorem 1.6. *If* $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_r}$ *are vectors in a vector space* V*, then:*

(a) The set W *of all linear combinations of* $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_r}$ *is a subspace of* V*.*

(b) W *is the smallest subspace of* V *that contains* $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_r}$ *every other subspace of* V *that contains* $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_r}$ *must contain* W

Proof. (a) To show that W is a subspace of V, it must be proven that it is closed under addition and scalar multiplication. There is at least one vector in W, namely, $\mathbf{0}$, since $\mathbf{0} = 0\mathbf{v_1} + 0\mathbf{v_2} + \cdots + 0\mathbf{v_r}$. If $\mathbf{u}$ and $\mathbf{v}$ are vectors in W, then

$$\mathbf{u} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_r\mathbf{v_r}$$

and

$$\mathbf{v} = k_1\mathbf{v_1} + k_2\mathbf{v_2} + \cdots + k_r\mathbf{v_r}$$

where $c_1, c_2, \dots, c_r, k_1, k_2, \dots, k_r$ are scalars. Therefore

$$\mathbf{u} + \mathbf{v} = (c_1 + k_1)\mathbf{v_1} + (c_2 + k_2)\mathbf{v_2} + \cdots + (c_r + k_r)\mathbf{v_r}$$

and, for any scalar k,

$$k\mathbf{u} = (kc_1)\mathbf{v_1} + (kc_2)\mathbf{v_2} + \cdots + (kc_r)\mathbf{v_r}$$

Thus, $\mathbf{u} + \mathbf{v}$ and $k\mathbf{u}$ are linear combinations of $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_r}$ and consequently lie in W. Therefore, W is closed under addition and scalar multiplication.

(b) Each vector $\mathbf{v_i}$ is a linear combination of $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}$ since we can write

$$\mathbf{v_i} = 0\mathbf{v_1} + 0\mathbf{v_2} + \cdots + 1\mathbf{v_i} + \cdots + 0\mathbf{v_r}$$

Therefore, the subspace W contains each of the vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}$. Let W' be any other subspace that contains $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}$. Since W' is closed under addition and scalar multiplication, it must contain all linear combinations of $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}$. Thus W' contains each vector of W.

$\square$

Definitions

- If S $= \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}\}$ is a set of vectors in a vector space V, then the subspace W of V consisting of all linear combinations of the vectors in S is called the **space spanned** by $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}$, and it is said that the vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}$ **span** W. To indicate that W is the space spanned by the vectors in the set $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}\}$ the below notation is used.

$$W = span(S) \quad or \quad W = span\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}\}$$

Examples The polynomials $1, x, x^2, \ldots, x^n$ span the vector space P_n defined previously since each polynomial $\mathbf{p}$ in P_n can be written as

$$\mathbf{p} = a_0 + a_1 x + \cdots + a_n x^n$$

which is a linear combination of $1, x, x^2, \ldots, x^n$. This can be denoted by writing

$$P_n = span\{1, x, x^2, \ldots, x^n\}$$

Spanning sets are not unique. For example, any two noncolinear vectors that lie in the $x - y$ plane will span the $x - y$ plane. Also, any nonzero vector on a line will span the same line.

Theorem 1.7. *Let* $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}\}$ *and* $S' = \{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_k}\}$ *be two sets of vectors in a vector space* V*. Then*

$$span(S) = span(S')$$

if and only if each vector in S *is a linear combination of those in* S' *and (conversely) each vector in* S' *is a linear combination of those in* S*.*

Proof. If each vector in S is a linear combination of those in S' then

$$span(S) \subseteq span(S')$$

and if each vector in S' is a linear combination of those in S then

$$span(S') \subseteq span(S)$$

and therefore

$$span(S) = span(S').$$

If

$$\mathbf{v_i} \neq a_1\mathbf{w_1} + a_2\mathbf{w_2} + \cdots + a_n\mathbf{w_n}$$

for all possible $a_1, a_2, \ldots, a_n$ then

$$\mathbf{v_i} \in span(S) \quad but \quad \mathbf{v_i} \notin span(S')$$

therefore

$$span(S) \neq span(S')$$

and vice versa. $\square$

1.4.12 Quiz

True or false?

(a) $\mathbf{0}$ is a linear combination of any non-empty set of vectors.

(b) If $S \subseteq V$ (vector space V), then $span(S)$ equals the intersection of all subspaces of V that contain S.

1.5 Linear Independence

In the previous section it was stated that a set of vectors S spans a given vector space V if every vector in V is expressible as a linear combination of the vectors in S. In general, it is possible that there may be more than one way to express a vector in V as a linear combination of vectors in a spanning set. This section will focus on the conditions under which each vector in V is expressible as a unique linear combination of the spanning vectors. Spanning sets with this property play a fundamental role in the study of vector spaces.

Definitions If $S = \{v_1, v_2, \ldots, v_r\}$ is a nonempty set of vectors, then the vector equation

$$k_1\mathbf{v_1} + k_2\mathbf{v_2} + \cdots + k_r\mathbf{v_r} = \mathbf{0}$$

has at least one solution, namely

$$k_1 = 0, k_2 = 0, \ldots, k_r = 0$$

If this is the only solution, then S is called a **linearly independent** set. If there are other solutions, then S is called a **linearly dependent** set.

Examples

1. If $\mathbf{v_1} = (2, -1, 0, 3), v_2 = (1, 2, 5, -1)$ and $v_3 = (7, -1, 5, 8)$, then the set of vectors $S = \{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$ is linearly dependent, since $3\mathbf{v_1} + \mathbf{v_2} - \mathbf{v_3} = \mathbf{0}$.

2. The polynomials

$$\mathbf{p_1} = 1 - x,\ \mathbf{p_2} = 5 + 3x - 2x^2,\ \mathbf{p_3} = 1 + 3x - x^2$$

form a linearly dependent set in P_2 since $3\mathbf{p_1} - \mathbf{p_2} + 2\mathbf{p_3} = \mathbf{0}$

3. Consider the vectors $\mathbf{i} = (1, 0, 0), \mathbf{j} = (0, 1, 0), \mathbf{k} = (0, 0, 1)$ in $\mathbb{R}^3$. In terms of components the vector equation

$$k_1\mathbf{i} + k_2\mathbf{j} + k_3\mathbf{k} = \mathbf{0}$$

becomes

$$k_1(1,0,0)+k_2(0,1,0)+k_3(0,0,1)=(0,0,0)$$

or equivalently,

$$(k_1,k_2,k_3)=(0,0,0)$$

Thus the set $S=\{\mathbf{i},\mathbf{j},\mathbf{k}\}$ is linearly independent. A similar argument can be used to extend S to a linear independent set in $\mathbb{R}^n$.

4. In $M_{2\times 3}(\mathbb{R})$, the set

$$\left\{\begin{pmatrix} 1 & -3 & 2 \\ -4 & 0 & 5 \end{pmatrix}, \begin{pmatrix} -3 & 7 & 4 \\ 6 & -2 & -7 \end{pmatrix}, \begin{pmatrix} -2 & 3 & 11 \\ -1 & -3 & 2 \end{pmatrix}\right\}$$

is linearly dependent since

$$5\begin{pmatrix} 1 & -3 & 2 \\ -4 & 0 & 5 \end{pmatrix}+3\begin{pmatrix} -3 & 7 & 4 \\ 6 & -2 & -7 \end{pmatrix}-2\begin{pmatrix} -2 & 3 & 11 \\ -1 & -3 & 2 \end{pmatrix}=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The following two theorems follow quite simply from the definition of linear independence and linear dependence.

Theorem 1.8. *A set S with two or more vectors is:*

(a) Linearly dependent if and only if at least one of the vectors in S is expressible as a linear combination of the other vectors in S.

(b) Linearly independent if and only if no vector in S is expressible as a linear combination of the other vectors in S.

Example

1. Recall that the vectors

$$\mathbf{v_1}=(2,-1,0,3),\ \mathbf{v_2}=(1,2,5,-1),\ \mathbf{v_3}=(7,-1,5,8)$$

were linear dependent because

$$3\mathbf{v_1} + \mathbf{v_2} - \mathbf{v_3} = \mathbf{0}.$$

It is obvious from the equation that

$$\mathbf{v_1} = \frac{-1}{3}\mathbf{v_2} + \frac{1}{3}\mathbf{v_3}, \quad \mathbf{v_2} = -3\mathbf{v_1} + 1\mathbf{v_3}, \quad \mathbf{v_3} = 3\mathbf{v_1} + 1\mathbf{v_2}$$

Theorem 1.9. *(a) A finite set of vectors that contains the zero vector is linearly dependent.*

(b) A set with exactly two vectors is linearly independent if and only if neither vector is a scalar multiple of the other.

2 BASIS AND DIMENSION

A line is thought of as 1-Dimensional, a plane 2-Dimensional, and surrounding space as 3-Dimensional. This section will attempt to make this intuitive notion of dimension precise and extend it to general vector spaces.

2.1 Coordinate systems of General Vector Spaces

A line is thought of as 1-Dimensional because every point on that line can be specified by 1 coordinate. In the same way a plane is thought of as 2 Dimensional because every point on that plane can be specified by 2 coordinates and so on. What defines this coordinate system? The most common form of defining a coordinate system is the use of coordinate axes. In the case of the plane the x and y axes are used most frequently. But there is also a way of specifying the coordinate system with vectors. This can be done by replacing each axis with a vector of length one that points in the positive direction of the axis. In the case of the $x - y$ plane the x and y-axes are replaced by the well known unit vectors $\mathbf{i}$ and $\mathbf{j}$ respectively. Let O be the origin of the system and P be any point in the plane. The point P can be specified by the vector $\overline{OP}$. Every vector, $\overline{OP}$ can be written as a linear combination of $\mathbf{i}$ and $\mathbf{j}$:

$$\overline{OP} = a\mathbf{i} + b\mathbf{j}$$

The coordinates of P, corresponding to this coordinate system, are (a, b).

Informally stated, vectors such as $\mathbf{i}$ and $\mathbf{j}$ that specify a coordinate system are called “basis vectors” for that system. Although in the preceding discussion our basis vectors were chosen to be of unit length and mutually perpendicular this is not essential. As long as linear combinations of the vectors chosen are capable of specifiying all points in the plane. In our example this only requires that the two vectors are not colinear. Different basis vectors however do change the coordinates of a point, as the following example demonstrates.

Example Let $S = \{\mathbf{i}, \mathbf{j}\}$, $U = \{\mathbf{i}, 2\mathbf{j}\}$ and $V = \{\mathbf{i} + \mathbf{j}, \mathbf{j}\}$. Let the sets S,U and V be three sets of basis vectors. Let P be the point $\mathbf{i} + 2\mathbf{j}$. The coordinates of P relative to each set of basis vectors is:

$$S \rightarrow (1, 2)$$

$$U \rightarrow (1, 1)$$

$$T \rightarrow (1, 1)$$

The following definition makes the preceding ideas more precise and enables the extension of a coordinate system to general vector spaces.

Definition

- If V is any vector space and $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is a set of vectors in V, then S is called a **basis** for V if the following two conditions hold:
  (a) S is linearly independent
  (b) S spans V

A basis is the vector space generalization of a coordinate system in 2-space and 3-space. The following theorem will aid in understanding how this is so.

Theorem 2.1. *If $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is a basis for a vector space V, then every vector $\mathbf{v}$ in V can be expressed in the form $\mathbf{v} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_n\mathbf{v_n}$ in exactly one way.*

Proof. Since S spans V, it follows from the definition of a spanning set that every vector in V is expressible as a linear combination of the vectors in S. To see that there is only one way to express a vector as a linear combination of the vectors in S, suppose that some vector $\mathbf{v}$ can be written as

$$\mathbf{v} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_n\mathbf{v_n}$$

and also as

$$\mathbf{v} = k_1\mathbf{v_1} + k_2\mathbf{v_2} + \cdots + k_n\mathbf{v_n}$$

Subtracting the second equation from the first gives

$$\mathbf{0} = (c_1 - k_1)\mathbf{v_1} + (c_2 - k_2)\mathbf{v_2} + \cdots + (c_n - k_n)\mathbf{v_n}$$

Since the right side of this equation is a linear combination of vectors in S, the linear independence of S implies that

$$(c_1 - k_1) = 0, (c_2 - k_2) = 0, \ldots, (c_n - k_n)$$

That is

$$c_1 = k_1, c_2 = k_2, \ldots, c_n = k_n$$

Thus the two expressions for $\mathbf{v}$ are the same. $\square$

Definitions

- If $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is a basis for a vector space V, and

  $$\mathbf{v} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_n\mathbf{v_n}$$

  is the expression for a vector $\mathbf{v}$ in terms of the basis S, then the scalars $c_1, c_2, \ldots, c_n$ are called the **coordinates** of $\mathbf{v}$ relative to the basis S. The vector $(c_1, c_2, \ldots, c_n)$ in $\mathbb{F}^n$ constructed from these coordinates is called the **coordinate vector of** v **relative to** S; it is denoted by

  $$[\mathbf{v}]_S = (c_1, c_2, \ldots, c_n)$$

- If $\mathbf{v} = [\mathbf{v}]_S$ then S is called the **standard basis**.

Remark It should be noted that coordinate vectors depend not only on the basis S but also on the order in which the basis vectors are written; a change in the order of the basis vectors results in a corresponding change of order for the entries in the coordinate vectors.

Examples

1. In example 3 of Section 1.5 it was shown that if

$$\mathbf{i} = (1,0,0), \ \ \mathbf{j} = (0,1,0), \ \ \mathbf{k} = (0,0,1)$$

then $S = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ is a linearly independent set in $\mathbb{R}^3$. This set also spans $\mathbb{R}^3$ since any vector $\mathbf{v} = (a,b,c)$ can be written as

$$\mathbf{v} = (a,b,c) = a(1,0,0) + b(0,1,0) + c(0,1,1) = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$$

Thus, S is a basis for $\mathbb{R}^3$. It is in fact a **standard basis** for $\mathbb{R}^3$. Looking at the coefficients of $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$ above, it follows that the coordinates of $\mathbf{v}$ relative to the standard basis are a, b and c, so

$$[\mathbf{v}]_S = (a,b,c)$$

and so we have

$$[\mathbf{v}]_S = \mathbf{v}.$$

2.2 Dimension of General Vector Spaces

Definition

- A nonzero vector space V is called **finite-dimensional** if it contains a finite set of vectors $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ that forms a basis. If no such set exists, V is called **infinite-dimensional**. In addition, the zero vector space is regarded as finite-dimensional.

Examples

- The vector spaces $\mathbb{F}^n$ and P_n are both finite-dimensional.

- The vector space of all real valued functions defined on $(-\infty, \infty)$ is infinite-dimensional.

Theorem 2.2. *If V is a finite-dimensional vector space and $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is any basis, then:*

(a) Every set with more than n vectors is linearly dependent.

(b) No set with fewer than n vectors spans V.

Proof. (a) Let $S' = \{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_m}\}$ be any set of m vectors in V, where $m > n$. It remains to be shown that S' is linearly dependent. Since $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is a basis for V, each $\mathbf{w_i}$ can be expressed as a linear combination of the vectors in S, say:

$$\mathbf{w_1} = a_{11}\mathbf{v_1} + a_{21}\mathbf{v_2} + \cdots + a_{n1}\mathbf{v_n}$$

$$\mathbf{w_2} = a_{12}\mathbf{v_1} + a_{22}\mathbf{v_2} + \cdots + a_{n2}\mathbf{v_n}$$

$$\vdots$$

$$\mathbf{w_m} = a_{1m}\mathbf{v_1} + a_{2m}\mathbf{v_2} + \cdots + a_{nm}\mathbf{v_n}$$

To show that S' is linearly dependent, scalars $k_1, k_2, \ldots, k_n$ must be found, not all zero, such that

$$k_1\mathbf{w_1} + k_2\mathbf{w_2} + \cdots + k_m\mathbf{w_m} = \mathbf{0}$$

combining the above 2 systems of equations gives

$$\begin{aligned}(k_1a_{11} + k_2a_{12} + \cdots + k_ma_{1m})\mathbf{v_1}& \\ + (k_1a_{21} + k_2a_{22} + \cdots + k_ma_{2m})\mathbf{v_2}& \\ \ddots& \\ + (k_1a_{n1} + k_2a_{n2} + \cdots + k_ma_{nm})\mathbf{v_n} &= \mathbf{0}\end{aligned}$$

Thus, from the linear independence of S, the problem of proving that S' is a linearly dependent set reduces to showing there are scalars $k_1, k_2, \ldots, k_m$, not all zero, that satisfy

$$a_{11}k_1 + a_{12}k_2 + \cdots + a_{1m}k_m = 0$$

$$a_{21}k_1 + a_{22}k_2 + \cdots + a_{2m}k_m = 0$$

$$\vdots$$

$$a_{n1}k_1 + a_{n2}k_2 + \cdots + a_{nm}k_m = 0$$

As the system is homogenous and there are more unknowns than equations ($m > n$), we have an infinite number of solutions, or in other words there are non trivial solutions such that $k_1, k_2, \ldots, k_m$ are not all zero.

(b) Let $S' = \{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_m}\}$ be any set of m vectors in V, where $m < n$. It remains to be shown that S' does not span V. The proof is by contradiction: assume S' spans V. This leads to a contradiction of the linear dependence of the basis $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ of V.

If S' spans V, then every vector in V is a linear combination of the vectors in S'. In particular, each basis vector $\mathbf{v_i}$ is a linear combination of the vectors in S', say

$$\mathbf{v_1} = a_{11}\mathbf{w_1} + a_{21}\mathbf{w_2} + \cdots + a_{n1}\mathbf{w_m}$$

$$\mathbf{v_2} = a_{12}\mathbf{w_1} + a_{22}\mathbf{w_2} + \cdots + a_{n2}\mathbf{w_m}$$

$$\vdots$$

$$\mathbf{v_n} = a_{1n}\mathbf{w_1} + a_{2n}\mathbf{w_2} + \cdots + a_{mn}\mathbf{w_m}$$

To obtain the contradiction it will be shown that there exist scalars $k_1, k_2, \ldots, k_n$ not all zero, such that

$$k_1\mathbf{v_1} + k_2\mathbf{v_2} + \cdots + k_n\mathbf{v_n} = \mathbf{0}$$

Observe the similarity to the above two systems compared with those given in the proof of (a). It can be seen that they are identical except that the $\mathbf{w}$'s and the $\mathbf{v}$'s and the m's and n's have been interchanged. Thus the above system in the same way again reduces to the problem of finding $k_1, k_2, \ldots, k_n$ not all zero, that satisfy

$$a_{11}k_1 + a_{12}k_2 + \cdots + a_{1m}k_n = 0$$

$$a_{21}k_1 + a_{22}k_2 + \cdots + a_{2m}k_n = 0$$

$$\vdots$$

$$a_{m1}k_1 + a_{m2}k_2 + \cdots + a_{mn}k_n = 0$$

As the system is homogenous and there are more unknowns than equations ($n > m$), we have an infinite number of solutions, or in other words there exist non trivial solutions such that $k_1, k_2, \ldots, k_m$ are not all zero. Hence the contradiction.

□

The last theorem essentially states the following. Let S be a set with n vectors which forms a basis for the vector space V. Let S' be another set of vectors in V consisting of m vectors. If m is greater than n, S' cannot form a basis for V as the vectors in S' cannot be linearly independent. If m is less than n, S' cannot form a basis for V because it does not span V. Thus, theorem 2.2 leads directly into one of the most important theorems in linear algebra.

Theorem 2.3. *All bases for a finite-dimenstional vector space have the same number of vectors.*

And thus the concept of dimension is almost complete. All that is needed is a definition.

Definition

- The **dimension** of a finite-dimensional vector space V, denoted by $dim(V)$, is defined to be the number of vectors in a basis for V. In addition, the zero vector space has dimension zero.

Examples

1. The dimensions of some common vector spaces are given below:

$$dim(\mathbb{F}^n) = n$$
$$dim(P_n) = n + 1$$
$$dim(M_{n\times n}(\mathbb{F})) = mn$$

2. Determine a basis (and hence dimension) for the solution space of the homogenous system:

$$\begin{array}{rcrcrcrcrcl} 2x_1 & + & 2x_2 & - & x_3 & & & + & x_5 & = & 0 \\ -x_1 & - & x_2 & + & 2x_3 & - & 3x_4 & + & x_5 & = & 0 \\ x_1 & + & x_2 & - & 2x_3 & & & - & x_5 & = & 0 \\ & & & & x_3 & + & x_4 & + & x_5 & = & 0 \end{array}$$

2.3 Related Theorems

The remaining part of this section states theorems which illustrate the subtle relationships among the concepts of spanning, linear independence, basis and dimension. In many ways these theorems form the building blocks of other results in linear algebra.

Theorem 2.4. Plus/Minus Theorem. *Let S be a nonempty set of vectors in a vector space V.*

(a) If S is a linearly independent set, and if $\mathbf{v}$ is a vector in V that is outside of the span(S), then the set $S \cup \{\mathbf{v}\}$ that results by inserting $\mathbf{v}$ is still linearly independent.

(b) If $\mathbf{v}$ is a vector in S that is expressible as a linear combination of other vectors in S, and if $S - \{\mathbf{v}\}$ denotes the set obtained by removing $\mathbf{v}$ from S, then S and $S - \{\mathbf{v}\}$ span the same space: that is,

$$span(S) = span(S - \{\mathbf{v}\})$$

A proof will not be included, but the theorem can be visualised in $\mathbb{R}^3$ as follows.

(a) Consider two linearly independent vectors in $\mathbb{R}^3$. These two vectors span a plane. If you add a third vector to them that is not in the plane, then the three vectors are still linearly independent and they span the entire domain of $\mathbb{R}^3$.

(b) Consider three non-colinear vectors in a plane that form a set S. The set S spans the plane. If any one of the vectors is removed from S to give S' it is clear that S' still spans the plane. That is $span(S) = span(S')$.

Theorem 2.5. *If V is an n-dimensional vector space and if S is a set in V with exactly n vectors, then S is a basis for V if either S spans V or S is linearly independent.*

Proof. Assume that S has exactly n vectors and spans V. To prove that S is a basis it must be shown that S is a linearly independent set. But if this is not so, then some vector $\mathbf{v}$ in S is a linear combination of the remaining vectors. If this vector is removed from S, then it follows from the theorem 2.4(b) that the remaining set of n-1 vectors still spans V. But this is impossible, since it follows from theorem 2.2(b), that no set with fewer than n vectors can span an n-dimensional vector space. Thus, S is linearly independent.

Assume S has exactly n vectors and is a linearly independent set. To prove that S is a basis it must be shown that S spans V. But if this is not so, then there is some vector $\mathbf{v}$ in V that is not in span(S). If this vector is inserted in S, then it follows from the theorem 2.4(a) that this set of n+1 vectors is still linearly independent. But this is impossible because it follows from theorem 2.2(a) that no set with more than n vectors in an n-dimensional vector space can be linearly independent. Thus S spans V. □

Examples

- $\mathbf{v_1} = (-3, 8)$ and $\mathbf{v_2} = (1, 1)$ form a basis for $\mathbb{R}^2$ because $\mathbb{R}^2$ has dimension two and v_1 and v_2 are linearly independent.

Theorem 2.6. *Let S be a finite set of vectors in a finite-dimensional vector space V.*

(a) If S spans V but is not a basis for V, then S can be reduced to a basis for V by removing appropriate vectors from S.

(b) If S is a linearly independent set that is not already a basis for V, then S can be enlarged to a basis for V by inserting appropriate vectors into S.

Proof. (a) The proof is constructive and is called the **left to right algorithm**.

Let $\mathbf{v_{c_1}}$ be the first nonzero vector in the set S. Choose the next vector in the list which is not a linear combination of $\mathbf{v_{c_1}}$ and call it $\mathbf{v_{c_2}}$. Find the next vector in the list which is not a linear combination of $\mathbf{v_{c_1}}$ and $\mathbf{v_{c_2}}$ and call it $\mathbf{v_{c_3}}$. Continue in such a way until the number of vectors chosen equals $dim(V)$.

(b) This proof is also constructive.

Let V be a vector space. Begin with $\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_r}$ which form a linearly independent family in V. Let $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ be a basis for V. Now it is necessary and important that $r < n$. To extend the basis, simply apply the left to right algorithm to the set (note that this set spans V because it contains a basis within it)

$$\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_r}, \mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$$

This will select a basis for V that commences with $\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_r}$

□

Theorem 2.7. *If W is a subspace of a finite-dimensional vector space V, then $dim(W) \leq dim(V)$; moreover, if $dim(W) = dim(V)$, then $W = V$*

Proof. Let $S = \{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_m}\}$ be a basis for W. Either S is also a basis for V or it is not. If it is, then $dim(W) = dim(V) = m$. If it is not, then by the previous theorem, vectors can be added to the linearly independent set S to make it into a basis for V, so $dim(W) < dim(V)$. Thus, $dim(W) \leq dim(V)$ in all cases. If $dim(W) = dim(V)$, then S is a set of m linearly independent vectors in the m-dimensional vector space V; hence by theorem 2.5, S is a basis for V. Therefore $W = V$. □

2.3.1 Quiz

True or false?

(a) The zero vector space has no basis.

(b) Every vector space that is spanned by a finite set has a basis.

(c) Every vector space has a finite basis.

(d) A vector space cannot have more than one basis.

(e) If a vector space has a finite basis, then the number of vectors in every basis is the same.

(f) Suppose that V is a finite dimensional vector space, S_1 is a linear independent subset of V, and S_2 is a subset of V that spans V. Then S_1 cannot contain more vectors than S_2.

(g) If S spans the vector space V, then every vector in V can be written as a linear combination of vectors in S in only one way.

(h) Every subspace of a finite dimensional vector space is finite dimensional.

(i) If V is an n dimensional vector space, then V has exactly one subspace with dimension 0 and one with dimension n.

(j) If V is an n dimensional vector space, and if S is a subset of V with n vectors, then S is linearly independent if and only if S spans V.

3 INNER PRODUCT SPACES AND ORTHONORMAL BASES

In many applications of vector spaces, we are concerned with the notion of measurement. In this section we introduce the idea of length through the structure of inner product spaces. We only consider $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

Definition

Let V be a vector space over $\mathbb{F}$. We define an inner product $\langle , \rangle$ on V to be a function that assigns a scalar $\langle \mathbf{u}, \mathbf{v} \rangle \in \mathbb{F}$ to every pair of ordered vectors $\mathbf{u}, \mathbf{v} \in V$ such that the following properties hold for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $\alpha \in \mathbb{F}$:

(a) $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$

(b) $\langle \alpha \mathbf{u}, \mathbf{v} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle$

(c) $\overline{\langle \mathbf{u}, \mathbf{v} \rangle} = \langle \mathbf{v}, \mathbf{u} \rangle$

(d) $\langle \mathbf{u}, \mathbf{u} \rangle \; > 0$ if $\mathbf{u} \neq \mathbf{0}$.

The main example is when $V = \mathbb{F}^n$. In this case we often use the notation $\langle \mathbf{u}, \mathbf{v} \rangle \equiv \mathbf{u} \cdot \mathbf{v}$ which is determined by

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} u_i \overline{v_i}$$

where $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$.

Definitions

- A vector space V over $\mathbb{F}$ endowed with a specific inner product is called an inner product space. If $\mathbb{F} = \mathbb{R}$ then V is said to be a real inner product space, whereas if $\mathbb{F} = \mathbb{C}$ we call V a complex inner product space.
- The norm (or length, or magnitude) of a vector $\mathbf{u}$ is given by $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$.

- Two vectors $\mathbf{u}, \mathbf{v}$ in an inner product space are said to be **orthogonal** if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

- If $\mathbf{u}$ and $\mathbf{v}$ are orthogonal vectors and both $\mathbf{u}$ and $\mathbf{v}$ have a magnitude of one (with respect to $\langle , \rangle$), then $\mathbf{u}$ and $\mathbf{v}$ are said to be **orthonormal**.

- A set of vectors in an inner product space is called an **orthogonal set** if all pairs of distinct vectors in the set are orthogonal. An orthogonal set in which each vector has a magnitude of one is called an **orthonormal set**.

The following additional properties follow easily from the axioms:

Theorem 3.1. *Let V be an inner product space, $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and $c \in \mathbb{F}$.*

(a) $\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle$.

(b) $\langle \mathbf{x}, c\mathbf{y} \rangle = \bar{c} \langle \mathbf{x}, \mathbf{y} \rangle$.

(c) $\langle \mathbf{x}, \mathbf{0} \rangle = \langle \mathbf{0}, \mathbf{x} \rangle = 0$.

(d) $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ *if and only if* $\mathbf{x} = \mathbf{0}$.

(e) If $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle$ *for all* $\mathbf{x} \in V$, *then* $\mathbf{y} = \mathbf{z}$.

Proof. (a) - (d) exercises

(e) By part (a) and (b), $\langle \mathbf{x}, \mathbf{y} - \mathbf{z} \rangle = 0$ for all $\mathbf{x} \in V$. Since this is true for all $\mathbf{x}$, it is true for $\mathbf{x} = \mathbf{y} - \mathbf{z}$, thus $\langle \mathbf{y} - \mathbf{z}, \mathbf{y} - \mathbf{z} \rangle = 0$. By (d) this implies that $\mathbf{y} = \mathbf{z}$. $\square$

Now that the groundwork has been laid the following theorem can be stated. The proof of this result is extremely important, since it makes use of an algorithm, or method, for converting an arbitary basis into an orthonormal basis.

Theorem 3.2. *Every non-zero finite dimensional inner product space V has an orthonormal basis.*

Proof. Let $\{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_m}\}$ be any basis for V. It suffices to show that V has an orthogonal basis, since the vectors in the orthogonal basis can be normalized to produce an orthonormal basis for V. The following sequence of steps will produce an orthogonal basis $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_m}\}$ for V.

Step 1 Let $\mathbf{v_1} = \mathbf{u_1}$.

Step 2 Obtain a vector $\mathbf{v_2}$ that is orthogonal to $\mathbf{v_1}$ by computing the component of $\mathbf{u_2}$ that is orthogonal to the space W_1 spanned by $\mathbf{v_1}$. This can be done using the formula:

$$\mathbf{v_2} = \mathbf{u_2} - \left(\frac{\langle \mathbf{u_2}, \mathbf{v_1} \rangle}{\langle \mathbf{v_1}, \mathbf{v_1} \rangle}\right) \mathbf{v_1}$$

Of course, if $\mathbf{v_2} = \mathbf{0}$, then $\mathbf{v_2}$ is not a basis vector. But this cannot happen, since it would then follow from the preceding formula for $\mathbf{v_2}$ that

$$\mathbf{u_2} = \left(\frac{\langle \mathbf{u_2}, \mathbf{v_1} \rangle}{\langle \mathbf{v_1}, \mathbf{v_1} \rangle}\right) \mathbf{v_1} = \left(\frac{\langle \mathbf{u_2}, \mathbf{v_1} \rangle}{\langle \mathbf{u_1}, \mathbf{u_1} \rangle}\right) \mathbf{u_1}$$

which says that $\mathbf{u_2}$ is a multiple of $\mathbf{u_1}$, contradicting the linear independence of the basis $S = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$.

Step 3 To construct a vector $\mathbf{v_3}$ that is orthogonal to both $\mathbf{v_1}$ and $\mathbf{v_2}$, compute the component of $\mathbf{u_3}$ orthogonal to the space W_2 spanned by $\mathbf{v_1}$ and $\mathbf{v_2}$ using the formula:

$$\mathbf{v_3} = \mathbf{u_3} - \left(\frac{\langle \mathbf{u_3}, \mathbf{v_1} \rangle}{\langle \mathbf{v_1}, \mathbf{v_1} \rangle}\right) \mathbf{v_1} - \left(\frac{\langle \mathbf{u_3}, \mathbf{v_2} \rangle}{\langle \mathbf{v_2}, \mathbf{v_2} \rangle}\right) \mathbf{v_2}$$

As in step 2, the linear independence of $\{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$ ensures that $\mathbf{v_3} \neq \mathbf{0}$. The remaining details are left as an exercise.

Step 4 To determine a vector $\mathbf{v_4}$ that is orthogonal to $\mathbf{v_1}, \mathbf{v_2}$ and $\mathbf{v_3}$, compute the component of $\mathbf{u_4}$ orthogonal to the space W_3 spanned by $\mathbf{v_1}, \mathbf{v_2}$ and $\mathbf{v_3}$ using the formula

$$\mathbf{v_4} = \mathbf{u_4} - \left(\frac{\langle \mathbf{u_4}, \mathbf{v_1} \rangle}{\langle \mathbf{v_1}, \mathbf{v_1} \rangle}\right) \mathbf{v_1} - \left(\frac{\langle \mathbf{u_4}, \mathbf{v_2} \rangle}{\langle \mathbf{v_2}, \mathbf{v_2} \rangle}\right) \mathbf{v_2} - \left(\frac{\langle \mathbf{u_4}, \mathbf{v_3} \rangle}{\langle \mathbf{v_3}, \mathbf{v_3} \rangle}\right) \mathbf{v_3}$$

Continuing in this way, an orthogonal set of vectors, $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_m}\}$, will be obtained after m steps. Since V is an m-dimensional vector space and every orthogonal set is linearly independent, the set $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is an orthogonal basis for V. $\square$

This preceding step-by-step construction for converting an arbitary basis into an orthogonal basis is called the **Gram-Schmidt process**.

Examples: *THE GRAM-SCHMIDT PROCESS*

1. Consider the vector space $\mathbb{R}^3$ with the Euclidean inner product. Apply the Gram-Schmidt process to transform the basis vectors $\mathbf{u_1} = (1,1,1), \mathbf{u_2} = (0,1,1), \mathbf{u_3} = (0,0,1)$ into an orthogonal basis $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$; then normalize the orthogonal basis vectors to obtain an orthonormal basis $\{\mathbf{q_1}, \mathbf{q_2}, \mathbf{q_3}\}$.

   Step 1

   $$\mathbf{v_1} = \mathbf{u_1} = (1,1,1)$$

   Step 2

   $$\begin{aligned} \mathbf{v_2} &= \mathbf{u_2} - \left(\frac{\mathbf{u_2}\cdot\mathbf{v_1}}{\mathbf{v_1}\cdot\mathbf{v_1}}\right)\mathbf{v_1} \\ &= (0,1,1) - \frac{2}{3}(1,1,1) = \left(\frac{-2}{3}, \frac{1}{3}, \frac{1}{3}\right) \end{aligned}$$

   Step 3

   $$\begin{aligned} \mathbf{v_3} &= \mathbf{u_3} - \left(\frac{\mathbf{u_3}\cdot\mathbf{v_1}}{\mathbf{v_1}\cdot\mathbf{v_1}}\right)\mathbf{v_1} - \left(\frac{\mathbf{u_3}\cdot\mathbf{v_2}}{\mathbf{v_2}\cdot\mathbf{v_2}}\right)\mathbf{v_2} \\ &= (0,0,1) - \frac{1}{3}(1,1,1) - \frac{1/3}{2/3}\left(\frac{-2}{3}, \frac{1}{3}, \frac{1}{3}\right) \\ &= \left(0, -\frac{1}{2}, \frac{1}{2}\right) \end{aligned}$$

   Thus,

   $$\mathbf{v_1} = (1,1,1), \quad \mathbf{v_2} = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right), \quad \mathbf{v_3} = \left(0, -\frac{1}{2}, \frac{1}{2}\right)$$

   form an orthogonal basis for $\mathbb{R}^3$. The norms of these vectors are

   $$\|\mathbf{v_1}\| = \sqrt{3}, \quad \|\mathbf{v_2}\| = \frac{\sqrt{6}}{3}, \quad \|\mathbf{v_3}\| = \frac{1}{\sqrt{2}}$$

   so an orthonormal basis for $\mathbb{R}^3$ is

   $$\mathbf{q_1} = \frac{\mathbf{v_1}}{\|\mathbf{v_1}\|} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \quad \mathbf{q_2} = \frac{\mathbf{v_2}}{\|\mathbf{v_2}\|} = \left(\frac{-2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$$

$$\mathbf{q_3} = \frac{\mathbf{v_3}}{\|\mathbf{v_3}\|} = \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

The Gramm-Schmidt process with subsequent normalization not only converts an arbitary basis $\{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$ into an orthonormal basis $\{\mathbf{q_1}, \mathbf{q_2}, \ldots, \mathbf{q_n}\}$, but it does it in such a way that for $k \geq 2$ the following relationships hold:

- $\{\mathbf{q_1}, \mathbf{q_2}, \ldots, \mathbf{q_k}\}$ is an orthonormal basis for the space spanned by $\{\mathbf{u_1}, \ldots, \mathbf{u_k}\}$.
- $\mathbf{q_k}$ is orthogonal to $\{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_{k-1}}\}$.

The proofs are omitted but these facts should become evident after some thoughtful examination of the proof of Theorem 3.1.

3.1 Quiz

True or false?

- An inner product is a scalar-valued function on the set of ordered pairs of vectors.
- An inner product space must be over the field of real or complex numbers.
- An inner product is linear in both components.
- If x, y and z are vectors in an inner product space such that $\langle x, y\rangle = \langle x, z\rangle$, then $y = z$.
- If $\langle x, y\rangle = 0$ for all x in an inner product space, then $y = 0$.

4 LINEAR TRANSFORMATIONS AND MATRICES

Definitions

- Let V, W be vector spaces over a field $\mathbb{F}$. A function that maps V into W, $T : V \to W$, is called a **linear transformation** from V to W if for all vectors $\mathbf{u}$ and $\mathbf{v}$ in V and all scalars $c \in \mathbb{F}$

  (a) $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$

  (b) $T(c\mathbf{u}) = cT(\mathbf{u})$

- In the special case where $V = W$, the linear transformation $T : V \to V$ is called a **linear operator** on V.

- Let A be an $m \times n$ matrix and let $T : \mathbb{F}^n \to \mathbb{F}^m$ be the linear transformation defined by $T(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^n$. Then as a matter of notational convention it is said that T is the linear transformation T_A.

4.0.1 Basic Properties of Linear Transformations

Theorem 4.1. *If $T : V \to W$ is a linear transformation, then:*

(a) If T is linear, then $T(\mathbf{0}) = \mathbf{0}$

(b) T is linear if and only if $T(a\mathbf{v} + \mathbf{w}) = aT(\mathbf{v}) + T(\mathbf{w})$ for all $\mathbf{v}, \mathbf{w}$ in V and $a \in \mathbb{F}$.

(c) $T(\mathbf{v} - \mathbf{w}) = T(\mathbf{v}) - T(\mathbf{w})$ for all $\mathbf{v}$ and $\mathbf{w}$ in V.

Part (a) of the above theorem states that a linear transformation maps $\mathbf{0}$ into $\mathbf{0}$. This property is useful for identifying transformations that are not linear. Part (b) is usually used to show that a transformation is linear.

Examples

1. T_A is a linear transformation. Let A be an $m \times n$ matrix and let $T : \mathbb{F}^n \rightarrow \mathbb{F}^m$ be the linear transformation defined by $T_A(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^n$. Let $\mathbf{u}$ and $\mathbf{v} \in \mathbb{F}^n$, then

   $$\begin{aligned} T(\lambda\mathbf{u} + \mathbf{v}) &= A(\lambda\mathbf{u} + \mathbf{v}) \\ &= \lambda A\mathbf{u} + A\mathbf{v} \\ &= \lambda T_A(\mathbf{u}) + T_A(\mathbf{v}) \end{aligned}$$

   and thus T_A is a linear transformation.

2. If I is the $n \times n$ identity matrix, then for every vector $\mathbf{x}$ in $\mathbb{F}^n$

   $$T_I(\mathbf{x}) = I\mathbf{x} = \mathbf{x}$$

   so multiplication by I maps every vector in $\mathbb{F}^n$ into itself. $T_I(\mathbf{x})$ is called the **identity operator** on $\mathbb{F}^n$.

3. Let A, B and X be $n \times n$ matrices. Then $Y = AX - XB$ is also $n \times n$.

   Let $V = M_{n\times n}(\mathbb{F})$ be the vector space of all $n\times n$ matrices. Then $Y = AX - XB$ defines a transformation $T : V \rightarrow V$. The transformation is linear since

   $$\begin{aligned} T(\lambda X_1 + X_2) &= A(\lambda X_1 + X_2) - (\lambda X_1 + X_2)B \\ &= \lambda AX_1 + AX_2 - \lambda X_1 b - X_2 B \\ &= \lambda(AX_1 - X_1 B) + AX_2 - X_2 B \\ &= \lambda T(X_1) + T(X_2) \end{aligned}$$

Theorem 4.2. *If $T : \mathbb{F}^n \rightarrow \mathbb{F}^m$ is a linear transformation, then there exists an $m \times n$ matrix A such that $T = T_A$.*

Example

1. Find the 2×2 matrix A such that $T = T_A$ has the property that

   $$T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \text{ and } T\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

4.1 Geometric Transformations in $\mathbb{R}^2$

This section consists of various different transformations of the form T_A that have a geometrical interpretation. Such transformations form the building blocks for understanding linear transformations.

Examples of Geometric Transformations

- Operators on $\mathbb{R}^2$ and $\mathbb{R}^3$ that map each vector into its symmetric image about some line or plane are called **reflection operators**. Such operators are of the form T_A and are thus linear. There are three main reflections in $\mathbb{R}^2$. These are summarised below. Considering the transformation from the coordinates (x, y) to (w_1, w_2) the properties of the operator are as follows.

  1. **Reflection about the y-axis**: The equations for this transformation are

     $$\begin{aligned} w_1 &= -x \\ w_2 &= y \end{aligned}$$

     The standard matrix for the transformation is clearly

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

     To demonstrate the reflection, consider the example below.

     $$\text{Let } \mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

     $$\text{therefore } T_A(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

  2. **Reflection about the x-axis**: The equations for this transformation are

     $$\begin{aligned} w_1 &= x \\ w_2 &= -y \end{aligned}$$

The standard matrix for the transformation is clearly

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

To demonstrate the reflection, consider the example below.

$$\text{Let } \mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\text{therefore } T_A(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

3. **Reflection about the line** $y = x$: The equations for this transformation are

$$\begin{aligned} w_1 &= y \\ w_2 &= x \end{aligned}$$

The standard matrix for the transformation is clearly

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

To demonstrate the reflection, consider the example below.

$$\text{Let } \mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\text{therefore } T_A(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

- Operators on $\mathbb{R}^2$ and $\mathbb{R}^3$ that map each vector into its orthogonal projection on a line or plane through the origin are called **orthogonal projection operators**.

Such operators are of the form T_A and are thus linear. There are two main projections in $\mathbb{R}^2$. These are summarised below. Considering the transformation from the coordinates (x, y) to (w_1, w_2) the properties of the operator are as follows.

1. **Orthogonal projection onto the x-axis**: The equations for this transformation are

   $$\begin{aligned} w_1 &= x \\ w_2 &= 0 \end{aligned}$$

   The standard matrix for the transformation is clearly

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

   To demonstrate the projection, consider the example below.

   $$\text{Let } \mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

   $$\text{therefore } T_A(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

2. **Orthogonal projection on the y-axis**: The equations for this transformation are

   $$\begin{aligned} w_1 &= 0 \\ w_2 &= y \end{aligned}$$

   The standard matrix for the transformation is clearly

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

To demonstrate the projection, consider the example below.

$$\text{Let } \mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\text{therefore } T_A(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

- An operator that rotates each vector in $\mathbb{R}^2$, through a fixed angle θ is called a **rotation operator** on $\mathbb{R}^2$. Such operators are of the form T_A and are thus linear. There is only one rotation in $\mathbb{R}^2$, due to the generality of the formula. This rotation is summarised below. Considering the transformation from the coordinates (x, y) to (w_1, w_2) the properties of the operator are as follows.

  1. **Rotation through an angle θ:** The equations for this transformation are

     $$\begin{aligned} w_1 &= x\cos\theta - y\sin\theta \\ w_2 &= x\sin\theta + y\cos\theta \end{aligned}$$

     The standard matrix for the transformation is clearly

     $$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

     To demonstrate the projection, consider the example below.

     $$\text{Let } \theta = 30^\circ \text{ and let } \mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

     $$\text{therefore } T_A(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}$$

- If k is a nonnegative scalar, then the operator $T(\mathbf{x}) = k\mathbf{x}$ on $\mathbb{R}^2$ and $\mathbb{R}^3$ is called a **contraction with factor** k if $0 \leq k \leq 1$, and a **dilation with factor** k if $k \geq 1$. Such operators are of the form T_A and are thus linear. The contraction and the dilation operators are summarised below. Considering the transformation from the coordinates (x, y) to (w_1, w_2) the properties of the operator are as follows.

  1. **Contraction with factor k on $\mathbb{R}^2$,** $(0 \leq k \leq 1)$: The equations for this transformation are.

     $$\begin{aligned} w_1 &= kx \\ w_2 &= ky \end{aligned}$$

     The standard matrix for the transformation is clearly

     $$A = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$$

     To demonstrate the contraction, consider the example below.

     $$\text{Let } k = \frac{1}{2} \text{ and let } \mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

     $$\text{therefore } T_A(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$$

  2. **Dilation with factor k on $\mathbb{R}^2$,** $(k \geq 1)$: The equations for this transformation are

     $$\begin{aligned} w_1 &= kx \\ w_2 &= ky \end{aligned}$$

     The standard matrix for the transformation is clearly

     $$A = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$$

To demonstrate the dilation, consider the example below.

$$\text{Let } k = 2 \text{ and let } \mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\text{therefore } T_A(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

4.2 Product of Linear Transformations

Definition

- If $T_1 : U \to V$ and $T_2 : V \to W$ are linear transformations, the **composite of T_2 with T_1** denoted by $T_2 \circ T_1$, is the function defined by the formula

  $$(T_2 \circ T_1)(\mathbf{u}) = T_2(T_1(\mathbf{u}))$$

  where $\mathbf{u}$ is a vector in U.

Remark: Observe that this definition requires the domain of T_2 (which is V) to contain the range of T_1; this is essential for the formula $T_2(T_1(\mathbf{u}))$ to make sense.

The next result shows that the composition of two linear transformations is itself a linear transformation.

Theorem 4.3. *If $T_1 : U \to V$ and $T_2 : V \to W$ are linear transformations, then $(T_2 \circ T_1) : U \to W$ is also a linear transformation.*

Proof. If $\mathbf{u}$ and $\mathbf{v}$ are vectors in U and $s \in \mathbb{F}$, then it follows from the definition of a composite transformation and from the linearity of T_1 and T_2 that

$$\begin{aligned} T_2 \circ T_1(s\mathbf{u} + \mathbf{v}) &= T_2(T_1(s\mathbf{u} + \mathbf{v})) \\ &= T_2(sT_1(\mathbf{u}) + T_1(\mathbf{v})) \\ &= sT_2(T_1(\mathbf{u})) + T_2(T_1(\mathbf{v})) \\ &= sT_2 \circ T_1(\mathbf{u}) + T_2 \circ T_1(\mathbf{v}) \end{aligned}$$

and thus the proof is complete. $\square$

Examples

1. Let A be an $m \times n$ matrix, and B be an $n \times p$ matrix, then AB is an $m \times p$ matrix. Also $T_A : \mathbb{F}^n \to \mathbb{F}^m$, and $T_B : \mathbb{F}^p \to \mathbb{F}^n$ are both linear transformations. Then

$$\begin{aligned} T_A \circ T_B &= T_A(T_B(\mathbf{x})) \\ &= AB\mathbf{x} \\ &= (AB)\mathbf{x} \\ &= T_{AB}(\mathbf{x}) \end{aligned}$$

   where $\mathbf{x} \in \mathbb{F}^p$. And therefore $T_A \circ T_B = T_{AB} : \mathbb{F}^p \to \mathbb{F}^m$.

2. If V has a basis $\beta = \{\mathbf{v_1}, \mathbf{v_2}\}$ and $T : V \to V$ is a linear transformation given by

$$\begin{aligned} T(\mathbf{v_1}) &= 2\mathbf{v_1} + 3\mathbf{v_2} \\ T(\mathbf{v_2}) &= -7\mathbf{v_1} + 8\mathbf{v_2} \end{aligned}$$

   To find $T \circ T(-\mathbf{v_1} + 3\mathbf{v_2})$ takes two steps as shown below.

$$\begin{aligned} T(-\mathbf{v_1} + 3\mathbf{v_2}) &= -T(\mathbf{v_1}) + 3T(\mathbf{v_2}) \\ &= -2\mathbf{v_1} - 3\mathbf{v_2} + 3(-7\mathbf{v_1} + 8\mathbf{v_2}) \\ &= -23\mathbf{v_1} + 21\mathbf{v_2} \end{aligned}$$

   Hence

$$\begin{aligned} T \circ T(-\mathbf{v_1} + 3\mathbf{v_2}) &= T(-23\mathbf{v_1} + 21\mathbf{v_2}) \\ &= -23T(\mathbf{v_1}) + 21T(\mathbf{v_2}) \\ &= -23(2\mathbf{v_1} + 3\mathbf{v_2}) + 21(-7\mathbf{v_1} + 8\mathbf{v_2}) \\ &= -193\mathbf{v_1} + 99\mathbf{v_2} \end{aligned}$$

4.3 Kernel and Image

Definitions

- If $T : V \to W$ is a linear transformation, then the set of vectors in V that T maps into $\mathbf{0}$ is called the **kernel** of T. It is denoted by $ker(T)$. In mathematical notation:
$$ker(T) = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0}\}$$

- If $T : V \to W$ is a linear transformation, then the set of all vectors in W that are images under T of at least one vector in V is called the **Image** (or range in some texts) of T; it is denoted by $Im(T)$. In mathematical notation:
$$Im(T) = \{\mathbf{w} \in W \mid \mathbf{w} = T(\mathbf{v}) \text{ for some } \mathbf{v} \in V\}$$

Examples

1. Let $I : V \to V$ be the identity operator. Since $I\mathbf{v} = \mathbf{v}$ for all vectors in V, every vector in V is the image of some vector (namely, itself); thus, $Im(I) = V$. Since the *only* vector that I maps into $\mathbf{0}$ is $\mathbf{0}$, it follows that $ker(I) = \{\mathbf{0}\}$.

2. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the orthogonal projection on the $x - y$ plane. The kernel of T is the set of points that T maps into $\mathbf{0} = (0, 0, 0)$; these are the points on the z-axis. Since T maps every point in $\mathbb{R}^3$ into the $x - y$ plane, the image of T must be some subset of this plane. But every point $(x_0, y_0, 0)$ in the $x - y$ plane is the image under T of some point; in fact, it is the image of all points on the vertical line that passes through $(x_0, y_0, 0)$. Thus $Im(T)$ is the entire $x - y$ plane.

3. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear operator that rotates each vector in the $x-y$ plane through the angle θ. Since *every* vector in the $x - y$ plane can be obtaine by rotating some vector through the angle θ, one obtains $Im(T) = \mathbb{R}^2$. Moreover, the only vector that rotates into $\mathbf{0}$ is $\mathbf{0}$, so $ker(T) = \{\mathbf{0}\}$.

4. Find the kernel and image of the linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ given by

$$T(\begin{bmatrix} x \\ y \\ z \end{bmatrix}) = \begin{bmatrix} x-y \\ 2z \end{bmatrix}$$

In all of the preceding examples, $ker(T)$ and $Im(T)$ turned out to be *subspaces*. This is no accident as the following theorem points out.

Theorem 4.4. *If $T : V \to W$ is a linear transformation, then:*

(a) The kernel of T is a subspace of V.

(b) The range of T is a subspace of W.

Proof. (a) To show that $ker(T)$ is a subspace, it must be shown that it contains at least one vector and is closed under addition and scalar multiplication. By part (a) of Theorem 4.1, the vector $\mathbf{0}$ is in $ker(T)$, so this set contains at least one vector. Let $\mathbf{v_1}$ and $\mathbf{v_2}$ be vectors in $ker(T)$, and let k be any scalar. Then

$$T(\mathbf{v_1} + \mathbf{v_2}) = T(\mathbf{v_1}) + T(\mathbf{v_2}) = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

so that $\mathbf{v_1}$ and $\mathbf{v_2}$ is in $ker(T)$. Also,

$$T(k\mathbf{v_1}) = kT(\mathbf{v_1}) = k\mathbf{0} = \mathbf{0}$$

so that $k\mathbf{v_1}$ is in $ker(T)$.

(b) Since $T(\mathbf{0}) = \mathbf{0}$, there is at least one vector in $Im(T)$. Let $\mathbf{w_1}$ and $\mathbf{w_2}$ be vectors in the range of T, and let k be any scalar. To prove this part it must be shown that $\mathbf{w_1} + \mathbf{w_2}$ and $k\mathbf{w_1}$ are in the range of T; that is, vectors $\mathbf{a}$ and $\mathbf{b}$ must be found in V such that $T(\mathbf{a}) = \mathbf{w_1} + \mathbf{w_2}$ and $T(\mathbf{b}) = k\mathbf{w_1}$.

Since $\mathbf{w_1}$ and $\mathbf{w_2}$ are in the range of T, there are vectors $\mathbf{a_1}$ and $\mathbf{a_2}$ in V such that $T(\mathbf{a_1}) = \mathbf{w_1}$ and $T(\mathbf{a_2}) = \mathbf{w_2}$. Let $\mathbf{a} = \mathbf{a_1} + \mathbf{a_2}$ and $\mathbf{b} = k\mathbf{a_1}$. Then

$$T(\mathbf{a}) = T(\mathbf{a_1} + \mathbf{a_2}) = T(\mathbf{a_1}) + T(\mathbf{a_2}) = \mathbf{w_1} + \mathbf{w_2}$$

and

$$T(\mathbf{b}) = T(k\mathbf{a_1}) = kT(\mathbf{a_1}) = k\mathbf{w_1}$$

which completes the proof.

□

Theorem 4.5. *If* $T : U \to V$ *is a linear transformation and* $\{\mathbf{u_1}, \mathbf{u_2}, \dots, \mathbf{u_n}\}$ *forms a basis for* U*, then* $Im(T) = span(T(\mathbf{u_1}), T(\mathbf{u_2}), \dots, T(\mathbf{u_n}))$

This theorem is best demonstrated by a simple example.

Example

Let A be $m \times n$ and let $T = T_A$. Then $T_A : \mathbb{F}^n \to \mathbb{F}^m$. Let $\{\mathbf{e_1}, \mathbf{e_2}, \ldots, \mathbf{e_n}\}$ be the standard basis for $\mathbb{F}^n$. Then by the previous theorem it can be stated

$$\begin{aligned} Im(T_A) &= span(T_A(\mathbf{e_1}), T_A(\mathbf{e_2}), \ldots, T_A(\mathbf{e_n})) \\ &= span(A\mathbf{e_1}, A\mathbf{e_2}, \ldots, A\mathbf{e_n}) \\ &= span(col_1(A), col_2(A), \ldots, col_n(A)) \end{aligned}$$

4.4 Rank and Nullity

Definitons If $T : U \to V$ is a linear transformation,

- the dimension of the image of T is called the **rank of** T and is denoted by rank(T),
- the dimension of the kernel is called the **nullity of** T and is denoted by nullity(T).

Example

- Let U be a vector space of dimension n, with basis $\{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$, and let $T : U \to U$ be a linear transformation defined by

  $$T(\mathbf{u_1}) = \mathbf{u_2}, T(\mathbf{u_2}) = \mathbf{u_3}, \cdots, T(\mathbf{u_{n-1}}) = \mathbf{u_n} \text{ and } T(\mathbf{u_n}) = \mathbf{0}$$

  Find bases for $ker(T)$ and $Im(T)$ and determine rank(T) and nullity(T).

Theorem 4.6. *If $T : U \to V$ is a linear transformation from an n-dimensional vector space U to a vector space V, then*

$$rank(T) + nullity(T) = dim(U) = n$$

Proof. The proof is divided up into two cases.

Case 1 Let U be the zero vector space. Then due to theorem 4.1 it is known that $T(\mathbf{0}) = \mathbf{0}$. Therefore it can be stated that

$$Im(T) = \{\mathbf{0}\} \text{ and } ker(T) = \{\mathbf{0}\}$$

therefore

$$\text{rank}(T)\text{+nullity}(T) = 0 + 0 = 0 = \dim(U)$$

Case 2 Let U be an n-dimensional vector space with the basis $\{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$. Then the proof can be divided into three parts.

(a) Consider the case where $ker(T) = \{\mathbf{0}\}$. Let $\mathbf{u} \in ker(T)$. As $\mathbf{u} \in U$ it can be expressed as

$$\mathbf{u} = x_1\mathbf{u_1} + x_2\mathbf{u_2} + \cdots + x_n\mathbf{u_n} \tag{1}$$

As $\mathbf{u} \in ker(T)$ it can be stated that

$$\mathbf{0} = T(\mathbf{u}) = x_1T(\mathbf{u_1}) + x_2T(\mathbf{u_2}) + \cdots + x_nT(\mathbf{u_n}) \tag{2}$$

Due to the fact that $ker(T) = \{\mathbf{0}\}$, $\mathbf{u} = \mathbf{0}$. Due to the linear independence of $\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}$ it follows from equation (1) that $x_1, x_2, \ldots, x_n = 0$. It then also follows from equation (2) that $T(\mathbf{u_1}), T(\mathbf{u_2}), \ldots, T(\mathbf{u_n})$ are linearly independent. It is known from Theorem 4.5 that $Im(T) = \text{span}(T(\mathbf{u_1}), T(\mathbf{u_2}), \ldots, T(\mathbf{u_n}))$. As $T(\mathbf{u_1}), T(\mathbf{u_2}), \ldots, T(\mathbf{u_n})$ are linearly independent they form a basis for $Im(T)$. It can therefore be stated that

$$\text{rank}(T)\text{+nullity}(T) = n + 0 = n = \dim(U)$$

(b) Consider the case where $ker(T) = U$. Theorem 4.5 states: $Im(T) = \text{span}(T(\mathbf{u_1}), T(\mathbf{u_2}), \ldots, T(\mathbf{u_n}))$. However $\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n} \in ker(T)$. Therefore $T(\mathbf{u_1}), T(\mathbf{u_2}), \ldots, T(\mathbf{u_n}) = \mathbf{0}$. So it can be stated that $Im(T) = \text{span}(\mathbf{0}) = \{\mathbf{0}\}$. Therefore

$$\text{rank}(T)\text{+nullity}(T) = 0 + n = n = \dim(U)$$

(c) Consider the case where $1 \leq \text{nullity}(T) < n$. Assume that the $\text{nullity}(T) = r$, and let $\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_r}$ be a basis for the kernel. Since $\{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_r}\}$ form a linearly independent set, theorem 2.6(b) states that there are $n - r$ vectors, $\mathbf{u_{r+1}}, \mathbf{u_{r+2}}, \ldots, \mathbf{u_n}$, such that $\{\mathbf{u_1}, \ldots, \mathbf{u_r}, \mathbf{u_{r+1}}, \ldots, \mathbf{u_n}\}$ is a basis for U. To complete the proof it shall be shown that the $n - r$ vectors in the set $S = \{T(\mathbf{u_{r+1}}), \ldots, T(\mathbf{u_n})\}$ form a basis for the image of T. It then follows that

$$\text{rank}(T)+\text{nullity}(T) = n - r + r = n = \dim(U)$$

First it shall be shown that S spans the image of T. If $\mathbf{b}$ is any vector in $Im(T)$, then $\mathbf{b} = T(\mathbf{u})$ for some vector $\mathbf{u}$ in U. Since $\{\mathbf{u_1}, \ldots, \mathbf{u_r}, \mathbf{u_{r+1}}, \ldots, \mathbf{u_n}\}$ is a basis for U, the vector $\mathbf{u}$ can be written in the form

$$\mathbf{u} = c_1\mathbf{u_1} + \cdots + c_r\mathbf{u_r} + c_{r+1}\mathbf{u_{r+1}} + \cdots + c_n\mathbf{u_n}$$

since $\mathbf{u_1}, \ldots, \mathbf{u_r}$ lie in the kernel of T, it is clear that $T(\mathbf{u_1}), \ldots, T(\mathbf{u_r}) = \mathbf{0}$, so that

$$\mathbf{b} = T(\mathbf{u}) = c_{r+1}T(\mathbf{u_{r+1}}) + \cdots + c_nT(\mathbf{u_n})$$

Thus, S spans the image of T.

Finally, it shall be shown that S is a linearly independent set and consequently forms a basis for $Im(T)$. Suppose that some linear combination of the vectors in S is zero; that is,

$$k_{r+1}T(\mathbf{u_{r+1}}) + \cdots + k_nT(\mathbf{u_n}) = \mathbf{0} \tag{3}$$

It must be shown that $k_{r+1} = \cdots = k_n = 0$. Since T is linear, equation (3) can be rewritten as

$$T(k_{r+1}\mathbf{u_{r+1}} + \cdots + k_n\mathbf{u_n}) = \mathbf{0}$$

which says that $k_{r+1}\mathbf{u_{r+1}} + \cdots + k_n\mathbf{u_n}$ is in the kernel of T. This vector can therefore be written as a linear combination of the basis vectors $\{\mathbf{u_1}, \ldots, \mathbf{u_r}\}$, say

$$k_{r+1}\mathbf{u_{r+1}} + \cdots + k_n\mathbf{u_n} = k_1\mathbf{u_1} + \cdots + k_r\mathbf{u_r}$$

Thus,

$$k_1\mathbf{u_1} + \cdots + k_r\mathbf{u_r} - k_{r+1}\mathbf{u_{r+1}} - \cdots - k_n\mathbf{u_n} = \mathbf{0}$$

Since $\{\mathbf{u_1}, \ldots, \mathbf{u_n}\}$ is linearly independent, all of the k's are zero; in particular $k_{r+1} = \cdots = k_n = 0$, which completes the proof.

□

Examples Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the linear operator that rotates each vector in the $x - y$ plane through an angle of θ. It was shown previously that $ker(T) = \{\mathbf{0}\}$ and $Im(T) = \mathbb{R}^2$. Thus,

$$\text{rank}(T)\text{+nullity}(T) = 2 + 0 = 2 = \dim(U)$$

which is consistent with the fact that the domain of T is two-dimensional.

4.5 Matrix of a Linear Transformation

In this section it shall be shown that if U and V are finite-dimensional vector spaces, then with a little ingenuity any linear transformation $T : U \rightarrow V$ can be regarded as a matrix transformation. The basic idea is to work with coordinate matrices of the vectors rather than with the vectors themselves.

Definition

- Suppose that U is an n-dimensional vector space and V an m-dimensional vector space. Let $T : U \rightarrow V$ be a linear transformation. Let β and γ be bases for U and V respectively, then for each $\mathbf{x}$ in U, the coordinate vector $[\mathbf{x}]_\beta$ will be a vector in $\mathbb{F}^n$, and the coordinate vector $[T(\mathbf{x})]_\gamma$ will be a vector in $\mathbb{F}^m$. If there exists an $m \times n$ matrix A, such that

$$A[\mathbf{x}]_\beta = [T(\mathbf{x})]_\gamma \tag{4}$$

then A is called **the matrix of the transformation relative to bases β and γ** and it is written

$$A = [T]_\beta^\gamma$$

Theorem 4.7. *Let $\beta = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$ and $\gamma = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_m}\}$ be bases for the vector spaces U and V respectively, and let $\mathbf{x} \in U$. If $T : U \to V$ is a linear transformation then*

(a) the matrix of the transformation relative to bases β and γ always exists. That is to say, there always exists a matrix $A = [T]_\beta^\gamma$ such that

$$A[\mathbf{x}]_\beta = [T(\mathbf{x})]_\gamma$$

(b) The matrix of the transformation relative to basis β and γ has the form

$$[T]_\beta^\gamma = [[T(\mathbf{u_1})]_\gamma \mid [T(\mathbf{u_2})]_\gamma \mid \cdots \mid [T(\mathbf{u_n})]_\gamma]$$

Proof. Let $\beta = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$ be a basis for the n-dimensional space U and let $\gamma = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_m}\}$ be a basis for the m-dimensional space V. Then the matrix $[T]_\beta^\gamma = A$ must have the form

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

such that (4) holds for all vectors $\mathbf{x}$ in U. In particular, this equation must hold for the basis vectors $\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}$; that is,

$$A[\mathbf{u_1}]_\beta = [T(\mathbf{u_1})]_\gamma, A[\mathbf{u_2}]_\beta = [T(\mathbf{u_2})]_\gamma, \ldots, A[\mathbf{u_n}]_\beta = [T(\mathbf{u_n})]_\gamma \tag{5}$$

But

$$[\mathbf{u_1}]_\beta = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, [\mathbf{u_2}]_\beta = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, [\mathbf{u_n}]_\beta = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

so

$$A[\mathbf{u_1}]_\beta = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}$$

$$A[\mathbf{u_2}]_\beta = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}$$

$$\vdots$$

$$A[\mathbf{u_n}]_\beta = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Substituting these results into equation (5) yields

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} = [T(\mathbf{u_1})]_\gamma, \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} = [T(\mathbf{u_2})]_\gamma, \ldots, \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = [T(\mathbf{u_n})]_\gamma$$

which shows that the successive columns of A are coordinate vectors of

$$T(\mathbf{u_1}), T(\mathbf{u_2}), \ldots, T(\mathbf{u_n})$$

with respect to the basis γ. Thus the matrix for T with respect to the bases β and γ is

$$[T]_\beta^\gamma = [[T(\mathbf{u_1})]_\gamma \mid [T(\mathbf{u_2})]_\gamma \mid \cdots \mid [T(\mathbf{u_n})]_\gamma]$$

Thus the proof is complete. □

Examples

1. Let $T_B : \mathbb{F}^n \to \mathbb{F}^m$ be a linear transformation defined by $T(X) = BX$ where B is an $m \times n$ matrix. Let $\beta = \{\mathbf{E_1}, \mathbf{E_2}, \ldots, \mathbf{E_n}\}$ be the standard basis for $\mathbb{F}^n$ and let $\gamma = \{\mathbf{e_1}, \mathbf{e_2}, \ldots, \mathbf{e_m}\}$ be the standard basis for $\mathbb{F}^m$. Then it is known from the previous theorem that $[T]_\beta^\gamma$ is the following matrix

   $$[T_B]_\beta^\gamma = [[T_B(\mathbf{E_1})]_\gamma \mid [T_B(\mathbf{E_2})]_\gamma \mid \cdots \mid [T_B(\mathbf{E_n})]_\gamma]$$

   In general, for $1 \leq j \leq n$ it follows from the definition of the transformation that

   $$T_B(\mathbf{E_j}) = B\mathbf{E_j} = col_j(B) = b_{1j}\mathbf{e_1} + b_{2j}\mathbf{e_2} + \cdots + b_{mj}\mathbf{e_m}$$

   therefore

   $$[T_B(\mathbf{E_j})]_\gamma = \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{mj} \end{bmatrix} = col_j(B)$$

and thus it is clear that

$$[T_B]_\beta^\gamma = B$$

2. Let U have the basis $\beta = \{\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}\}$ and let V have the basis $\gamma = \{\mathbf{v_1}, \mathbf{v_2}\}$. Let T be the linear transformation defined by

$$T(\mathbf{u_1}) = 2\mathbf{v_1} + \mathbf{v_2},\ T(\mathbf{u_2}) = \mathbf{v_1} - \mathbf{v_2},\ T(\mathbf{u_3}) = 2\mathbf{v_2}$$

Then clearly

$$[T]_\beta^\gamma = \begin{bmatrix} 2 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix}$$

3. Let $V = M_{2\times 2}(\mathbb{R})$ and let $T : V \to V$ be the linear transformation given by $T(X) = BX - XB$ where $X \in V$ and

$$B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Let $\beta = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ be the standard basis for V where

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\ E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\ E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},\ E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

To find $[T]_\beta^\beta$ it is necessary to do the following calculations:

$$T(E_{11}) = BE_{11} - E_{11}B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$= \begin{bmatrix} 0 & -b \\ c & 0 \end{bmatrix} = 0E_{11} + -bE_{12} + cE_{21} + 0E_{22}$$

$$T(E_{12}) = BE_{12} - E_{12}B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$= \begin{bmatrix} -c & a-d \\ 0 & c \end{bmatrix} = -cE_{11} + (a-d)E_{12} + 0E_{21} + cE_{22}$$

$$T(E_{21}) = BE_{21} - E_{21}B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$= \begin{bmatrix} b & 0 \\ d-a & -b \end{bmatrix} = bE_{11} + 0E_{12} + (d-a)E_{21} + -bE_{22}$$

$$T(E_{22}) = BE_{22} - E_{22}B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$= \begin{bmatrix} 0 & b \\ -c & 0 \end{bmatrix} = 0E_{11} + bE_{12} + -cE_{21} + 0E_{22}$$

Therefore it follows that

$$[T]_{\beta}^{\beta} = \begin{bmatrix} 0 & -c & b & 0 \\ -b & a-d & 0 & b \\ c & 0 & d-a & -c \\ 0 & c & -b & 0 \end{bmatrix}$$

The following theorem follows directly from the definition of the matrix of a linear transformation.

Theorem 4.8. *Let $T : U \to V$ be a linear transformation, and let β and γ be bases for U and V respectively. Then if $\mathbf{u} \in U$*

$$[T(\mathbf{u})]_{\gamma} = [T]_{\beta}^{\gamma}[\mathbf{u}]_{\beta}$$

Examples

1. Let U have the basis $\beta = \{\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}\}$ and let V have the basis $\gamma = \{\mathbf{v_1}, \mathbf{v_2}\}$. Let T be the linear transformation defined by

$$T(\mathbf{u_1}) = 2\mathbf{v_1} + \mathbf{v_2},\ T(\mathbf{u_2}) = \mathbf{v_1} - \mathbf{v_2},\ T(\mathbf{u_3}) = 2\mathbf{v_2}$$

Given that $\mathbf{u} = 3\mathbf{u_1} + -2\mathbf{u_2} + 7\mathbf{u_3}$

$$[\mathbf{u}]_{\beta} = \begin{bmatrix} 3 \\ -2 \\ 7 \end{bmatrix},\ \text{and}\ [T]_{\beta}^{\gamma} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix}$$

Hence

$$[T(\mathbf{u})]_\gamma = [T]_\beta^\gamma [\mathbf{u}]_\beta = \begin{bmatrix} 2 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ 7 \end{bmatrix} = \begin{bmatrix} 4 \\ 19 \end{bmatrix}$$

Hence $T(\mathbf{u}) = 4\mathbf{v_1} + 19\mathbf{v_2}$.

Example

1. Let $T : V \to W$ be a linear transformation, and let $\beta = \{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$ and $\gamma = \{\mathbf{w_1}, \mathbf{w_2}, \mathbf{w_3}\}$ be bases for V and W respectively. T is the linear transformation given by

$$\begin{aligned} T(\mathbf{v_1}) &= \mathbf{w_1} + \mathbf{w_2} + -\mathbf{w_3} \\ T(\mathbf{v_2}) &= 2\mathbf{w_1} + -3\mathbf{w_2} \\ T(\mathbf{v_3}) &= 3\mathbf{w_1} + -2\mathbf{w_2} + -\mathbf{w_3} \end{aligned}$$

Find $ker(T)$ and $Im(T)$.

The following theorem gives a recipe for finding bases for $ker(T)$ and the $Im(T)$ where possible.

Theorem 4.9. *Let A be an $m \times n$ matrix such that $A = [T]_\beta^\gamma$, where $T : V \to W$ is a linear transformation and $\beta = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ and $\gamma = \{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_m}\}$ are bases for V and W respectively. Let $s = nullity(A)$ and $r = rank(A)$. Then suppose that*

$$\mathbf{x_j} = \begin{bmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{nj} \end{bmatrix}$$

for $1 \leq j \leq s$, form a basis for $N(A)$, while $col_{c_1}(A), col_{c_2}(A), \ldots, col_{c_t}(A)$ form a basis for $C(A)$ Then

1. (a) *the vectors* $\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_s}$ *defined by*

   $$\mathbf{u_j} = x_{1j}\mathbf{v_1} + x_{2j}\mathbf{v_2} + \cdots + x_{nj}\mathbf{v_n}$$

   will be a basis for the kernel of T*.*

   (b) *the vectors* $T(\mathbf{v_{c_1}}), T(\mathbf{v_{c_2}}), \ldots, T(\mathbf{v_{c_r}})$ *form a basis for the image of* T*.*

2. *If* $N(A) = \{\mathbf{0}\}$*, then* $ker(T) = \{\mathbf{0}\}$

   If $C(A) = \{\mathbf{0}\}$*, then* $Im(T) = \{\mathbf{0}\}$

3. *Thus it follows that rank(*T*) = rank(*A*) and nullity(*T*) = nullity(*A*).*

Quiz: True or false?

Let U and V be vector spaces of dimension n and m respectively over a field $\mathbb{F}$, and let T be a linear transformation from U to V. Let $\beta = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be a basis for U and let $\gamma = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ be a basis for V. Let $\mathbf{u} \in U$.

(a) For any $a_1, \ldots, a_n \in \mathbb{F}$,

$$T(\sum_{i=1}^{n} a_i \mathbf{u}_i) = \sum_{i=1}^{n} a_i T(\mathbf{u}_i).$$

(b) $\{T(\mathbf{u}_1), \ldots, T(\mathbf{u}_n)\}$ is a basis for $\mathrm{Im}(T)$.

(c) If $nullity(T) = 0$ then $m = n$.

(d) $rank(T) + nullity(T) = n$

(e) If α is a basis for $\ker(T)$ and $\alpha \subseteq \beta$, then $\beta \setminus \alpha$ is a basis for $\mathrm{Im}(T)$.

(f) If $U = \mathbb{F}^n$, then $[\mathbf{u}]_\beta = \mathbf{u}$.

(g) $[\mathbf{u}_i]_\beta = \mathbf{e}_i$.

(h) $[\mathbf{u}]_\beta$ will be a column vector if and only if $U = \mathbb{F}^n$.

(i) $[\mathbf{u}]_\beta$ depends on the order of β.

(j) If $A = [T]_\beta^\gamma$, then $col_i A = [T(\mathbf{u}_i)]_\beta$.

(k) If $m = n$ and $\mathbf{u}_i = \mathbf{v}_i$ for all i, then $[T]_\beta^\gamma = I$.

(l) To get $[T]_\beta^\gamma$, we need to calculate $T(\mathbf{u}_i)$ for each i, and express the answer as a linear combination of vectors from γ.

4.6 T-invariant subspaces

Definition

Let $T : V \to V$ be a linear operator. A subspace W of V is called T-invariant if $T(w) \in W, \forall w \in W$.

Examples

The are many examples of T-invariant subspaces. Veriy that the following are all T-invariant:

- $\{0\}$
- V
- $ker(T)$
- $Im(T)$
- E_λ which is the space spanned by linearly independent eigenvectors of T corresponding to eigenvalue λ. (*)
- A T-cyclic subspace of V generated by $v \in V$, given by $span\{v, T(v), T^2(v), \ldots\}$

4.7 Vector space of linear transformations

If $f, g : V \to W$ are functions and V, W are vector spaces over $\mathbb{F}$ we have seen that we can define addition and scalar multiplication by $(f + g)(v) = f(v) + g(v)$ and $(af)(v) = af(v)$ with $a \in \mathbb{F}$ and $v \in V$.

Using the above definition, it is easily verified that if T_1, T_2 are linear transformations, then the linear combination $aT_1 + T_2$ is also a linear transformation. In fact, the set of all linear transformations from V to W is itself a vector space, denoted $\ell(V, W)$. In the case $V = W$ we often write $\ell(V)$.

In fact we have the relationships

$$\begin{aligned}[T_1+T_2]_\beta^\gamma &= [T_1]_\beta^\gamma+[T_2]_\beta^\gamma,\\ [aT]_\beta^\gamma &= a[T]_\beta^\gamma.\end{aligned}$$

This is leading up to the notion of associating the vector space $\ell(V,W)$ with $M_{m\times n}$ in the case V and W are of dimension n and m respectively. Before making this identification we should investigate the concept of isomorphic vector spaces.

4.8 Isomorphisms and inverses of transformations

A linear transformation $T:V\to W$ is said to be invertible if there exists a unique transformation $T^{-1}:W\to V$ such that $T\circ T^{-1}=I_W$ and $T^{-1}\circ T=I_V$. We call T^{-1} the inverse of T.

We have the following facts regarding inverses:

- A function is invertible if and only if it is one-to-one and onto.
- T^{-1} is linear.
- T is invertible if and only if $rank(T)=\dim(V)=\dim(W)$.
- $[T^{-1}]_\gamma^\beta=([T]_\beta^\gamma)^{-1}$.

We say that V is **isomorphic** to W if there exists an invertible linear transformation $T:V\to W$. We write $V\cong W$ to indicate that V is isomorphic to W. Such a T is called an isomorphism.

The main result of this section is the following:

If V and W are finite dimensional vector spaces over the same field, then $V\cong W$ if and only if $\dim(V)=\dim(W)$.

Examples

(a) $T:\mathbb{F}^2\to P_2(\mathbb{F})$ via $T(a,b)=a+bx$.

(b) $T:P_3(\mathbb{F})\to M_2(\mathbb{F})$ via $T(a+bx+cx^2+dx^3)=\begin{pmatrix} a+b & b+c\\ c+d & d\end{pmatrix}$.

A corollary to our result is that for a vector space V over $\mathbb{F}$, V is isomorphic to $\mathbb{F}^n$ if and only if $\dim(V) = n$.

This formalises our association of n dimensional vector spaces with $\mathbb{F}^n$ as I hinted at when we looked at standard bases.

Another consequence is that we can associate $\ell(V, W)$ with $M_{m\times n}$.

4.9 Change of Basis

A basis of a vector space is a set of vectors that specify the coordinate system. A vector space may have an infinite number of bases but each basis contains the same number of vectors. The number of vectors in the basis is called the dimension of the vector space. The coordinate vector or coordinate matrix of a point changes with any change in the basis used. If the basis for a vector space is changed from some old bases β to some new bases γ, how is the old coordinate vector $[\mathbf{v}]_\beta$ of a vector $\mathbf{v}$ related to the new coordinate vector $[\mathbf{v}]_\gamma$? The following theorem answers that question.

Theorem 4.10. *If the basis for a vector space is changed from some old basis* $\beta = \{\mathbf{u_1}, \mathbf{u_2}, \dots, \mathbf{u_n}\}$ *to some new basis* $\gamma = \{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$*, then the old coordinate vector* $[\mathbf{w}]_\beta$ *is related to the new coordinate vector* $[\mathbf{w}]_\gamma$ *of the same vector* $\mathbf{w}$ *by the equation*

$$[\mathbf{w}]_\gamma = P[\mathbf{w}]_\beta$$

where the columns of P *are the coordinate vectors of the old basis vectors relative to the new basis; that is, the column vectors of* P *are*

$$[\mathbf{u_1}]_\gamma, [\mathbf{u_2}]_\gamma, \dots, [\mathbf{u_n}]_\gamma$$

P is called the **change of basis matrix** *or the* **change of coordinate matrix***.*

Proof. Let V be a vector space with a basis $\beta = \{\mathbf{u_1}, \mathbf{u_2}, \dots, \mathbf{u_n}\}$ and a new basis

$\gamma = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$. Let $\mathbf{w} \in V$. Therefore $\mathbf{w}$ can be expressed as:

$$\mathbf{w} = a_1\mathbf{u_1} + a_2\mathbf{u_2} + \cdots + a_n\mathbf{u_n}$$

Thus we have

$$[\mathbf{w}]_\beta = (a_1, a_2, \ldots, a_n)$$

As γ is also a basis of V the elements of β can be expressed as follows

$$\mathbf{u_1} = p_{11}\mathbf{v_1} + p_{21}\mathbf{v_2} + \cdots + p_{n1}\mathbf{v_n}$$

$$\mathbf{u_2} = p_{12}\mathbf{v_1} + p_{22}\mathbf{v_2} + \cdots + p_{n2}\mathbf{v_n}$$

$$\vdots$$

$$\mathbf{u_n} = p_{1n}\mathbf{v_1} + p_{2n}\mathbf{v_2} + \cdots + p_{nn}\mathbf{v_n}$$

Combining this system of equations with the above expression for $\mathbf{w}$ gives

$$\mathbf{w} = (p_{11}a_1 + p_{12}a_2 + \cdots + p_{1n}a_n)\mathbf{v_1}$$

$$+(p_{21}a_1 + p_{22}a_2 + \cdots + p_{2n}a_n)\mathbf{v_2}+$$

$$\vdots$$

$$+(p_{n1}a_1 + p_{n2}a_2 + \cdots + p_{nn}a_n)\mathbf{v_n}+$$

and thus it can be seen that

$$[\mathbf{w}]_\gamma = \begin{bmatrix} p_{11}a_1 + p_{12}a_2 + \cdots + p_{1n}a_n \\ p_{21}a_1 + p_{22}a_2 + \cdots + p_{2n}a_n \\ \vdots \\ p_{n1}a_1 + p_{n2}a_2 + \cdots + p_{nn}a_n \end{bmatrix}$$

which can be written as

$$[\mathbf{w}]_\gamma = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

from which it can be seen

$$[\mathbf{w}]_\gamma = P[\mathbf{w}]_\beta$$

where P's columns are

$$[\mathbf{u_1}]_\gamma, [\mathbf{u_2}]_\gamma, \ldots, [\mathbf{u_n}]_\gamma$$

$\square$

Example

1. Consider the bases $\gamma = \{\mathbf{v_1}, \mathbf{v_2}\}$ and $\beta = \{\mathbf{u_1}, \mathbf{u_2}\}$ for $\mathbb{R}^2$, where

$$\mathbf{v_1} = (1, 0); \quad \mathbf{v_2} = (0, 1); \quad \mathbf{u_1} = (1, 1); \quad \mathbf{u_2} = (2, 1)$$

   (a) Find the transition matrix from β to γ. First the coordinate vectors of the old basis vectors $\mathbf{u_1}$ and $\mathbf{u_2}$ must be found relative to the new basis β. By inspection:

$$\mathbf{u_1} = \mathbf{v_1} + \mathbf{v_2}$$

$$\mathbf{u_2} = 2\mathbf{v_1} + \mathbf{v_2}$$

   so that

$$[\mathbf{u_1}]_\gamma = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \textit{and} \quad [\mathbf{u_2}]_\gamma = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

   Thus the transition matrix from β to γ

$$P = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$$

   (b) Use the transition matrix to find $[\mathbf{v}]_\gamma$ if

$$[\mathbf{v}]_\beta = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$$

   It is known from the above change of basis theorem 4.10 that

$$[\mathbf{v}]_\gamma = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ 5 \end{bmatrix} = \begin{bmatrix} 7 \\ 2 \end{bmatrix}$$

As a check it should be possible to recover the vector $\mathbf{v}$ either from $[\mathbf{v}]_\beta$ or $[\mathbf{v}]_\gamma$. It is left for the student to show that $-3\mathbf{u_1}+5\mathbf{u_2} = 7\mathbf{v_1}+2\mathbf{v_2} = (7,2)$.

4.10 Similar Matrices

The matrix of a linear operator $T : V \to V$ depends on the basis selected for V. One of the fundamental problems of linear algebra is to choose a basis for V that makes the matrix for T as simple as possible - diagonal or triangular, for example. This section is devoted to the study of this problem

To demonstrate that certain bases produce a much simpler matrix of transformation than others, consider the following example.

Example

1. Standard bases do not necessarily produce the simplest matrices for linear operators. For example, consider the linear operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ -2x_1 + 4x_2 \end{bmatrix}$$

and the standard basis $\beta = \{\mathbf{e_1}, \mathbf{e_2}\}$ for $\mathbb{R}^2$, where

$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

By theorem 4.7, the matrix for T with respect to this basis is the standard matrix for T; that is,

$$[T]_\beta^\beta = [T(\mathbf{e_1}) \mid T(\mathbf{e_2})]$$

From the definition of the linear transformation T,

$$T(\mathbf{e_1}) = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad T(\mathbf{e_2}) = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

so

$$[T]_\beta^\beta = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$$

In comparison, consider the basis $\gamma = \{\mathbf{u_1}, \mathbf{u_2}\}$, where

$$\mathbf{u_1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{u_2} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

By theorem 4.7, the matrix for T with respect to the basis γ is

$$[T]_\gamma^\gamma = [[T(\mathbf{u_1})]_\gamma \mid [T(\mathbf{u_2})]_\gamma$$

From the definiton of the linear transformation T,

$$T(\mathbf{u_1}) = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = 2\mathbf{u_1}, \quad T(\mathbf{u_2}) = \begin{bmatrix} 3 \\ 6 \end{bmatrix} = 3\mathbf{u_2}$$

Hence

$$[T(\mathbf{u_1})]_\gamma = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \quad [T(\mathbf{e_2})]_\gamma = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

So

$$[T]_\gamma^\gamma = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

This matrix is 'simpler' in the sense that diagonal matrices enjoy special properties that more general matrices do not.

Much research has been devoted to determining the “simplest possible form” that can be obtained for the matrix of a linear operator $T : V \to V$, by choosing the basis appropriately. This problem can be attacked by first finding a matrix for T relative to *any* basis, say a standard basis, where applicable, then changing the basis in a manner that simplifies the matrix. Before pursuing this idea further, it is necessary to grasp the theorem below. It gives a useful alternative viewpoint about change of basis matrices; it shows that the transition matrix form a basis β to γ can be regarded as the matrix of transformation of the identity operator.

Theorem 4.11. *If β and γ are bases for a finite-dimensional vector space V, and if $I : V \to V$ is the identity operator, then $[I]_\beta^\gamma$ is the transition matrix form β to γ.*

Proof. Suppose that $\beta = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$ and $\gamma = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ are bases for V. Using the fact that $I(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in V$, it follows that

$$\begin{aligned} [I]_\beta^\gamma &= [[I(\mathbf{u_1})]_\gamma \mid [I(\mathbf{u_2})]_\gamma \mid [I(\mathbf{u_n})]_\gamma] \\ &= [[\mathbf{u_1}]_\gamma \mid [\mathbf{u_2}]_\gamma \mid [\mathbf{u_n}]_\gamma] \end{aligned}$$

which is the change of basis matrix from β to γ. □

The ground work has been laid to consider the main problem in this section.

Problem: If β and γ are two bases for a finite-dimensional vector space V, and if $T : V \to V$ is a linear operator, what relationship, if any, exists between the matrices $[T]_\beta^\beta$ and $[T]_\gamma^\gamma$?

The answer to this question can be obtained by considering the composition of three linear operators. Consider a vector $\mathbf{v} \in V$. Let the vector $\mathbf{v}$ be mapped into itself by the identity operator, then let $\mathbf{v}$ be mapped into $T(\mathbf{v})$ by T, then let $T(\mathbf{v})$ be mapped into itself by the identity operator. All four vector spaces involved in the composition are the same (namely V); however, the bases for the spaces vary. Since the starting vector is $\mathbf{v}$ and the final vector is $T(\mathbf{v})$, the composition is the same as T; that is,

$$T = I \circ T \circ I$$

If the first and last vector spaces are assigned the basis γ and the middle two spaces are assigned the basis β, then it follows from the previous statement $T = I \circ T \circ I$, that

$$[T]_\gamma^\gamma = [I \circ T \circ I]_\gamma^\gamma = [I]_\beta^\gamma [T]_\beta^\beta [I]_\gamma^\beta$$

But $[I]_\gamma^\beta$ is the change of basis matrix from γ to β and consequently $[I]_\beta^\gamma$ is the change of basis matrix from β to γ. Thus, let $P = [I]_\gamma^\beta$, then $P^{-1} = [I]_\beta^\gamma$ and hence it can be written that

$$[T]_\gamma^\gamma = P^{-1}[T]_\beta^\beta P$$

This is all summarised in the following theorem.

Theorem 4.12. *Let $T : V \to V$ be a linear operator on a finite-dimensional vector space V, and let β and γ be bases for V. Then*

$$[T]_\gamma^\gamma = P^{-1}[T]_\beta^\beta P \tag{6}$$

where P is the change of basis matrix from γ to β.

Remark: When applying theorem 4.12, it is easy to forget whether P is the change of basis matrix from β to γ or the change of basis matrix from γ to β. Just remember that in order for $[T]_\beta^\beta$ to operate succesfully on a vector $\mathbf{v}$, $\mathbf{v}$ must be in the basis β. Therefore, due to P's positioning in the formula, it must be the change of basis matrix from γ to β.

Example

1. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ -2x_1 + 4x_2 \end{bmatrix}$$

Find the matrix of T with respect to the standard basis $\beta = \{\mathbf{e_1}, \mathbf{e_2}\}$ for $\mathbb{R}^2$, then use theorem 4.12 to find the matrix of T with respect to the basis $\gamma = \{\mathbf{u_1}, \mathbf{u_2}\}$, where

$$\mathbf{u_1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{u_2} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

It was shown earlier that

$$[T]_\beta^\beta = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$$

To find $[T]_\gamma^\gamma$ from (6), requires the change of basis matrix P, where

$$P = [I]_\gamma^\beta = [[\mathbf{u_1}]_\beta \mid [\mathbf{u_2}]_\beta]$$

By inspection

$$\begin{aligned} \mathbf{u_1} &= \mathbf{e_1} + \mathbf{e_2} \\ \mathbf{u_2} &= \mathbf{e_1} + 2\mathbf{e_2} \end{aligned}$$

so that

$$[\mathbf{u_1}]_\beta = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad [\mathbf{u_2}]_\beta = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Thus the transition matrix from γ to β is

$$P = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

It is clear that

$$P^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

so that by theorem 4.12 the matrix of T relative to the basis γ is

$$[T]_\gamma^\gamma = P^{-1}[T]_\beta^\beta P = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

which agrees with the previous result.

The relationship in (6) is of such importance that there is some terminology associated with it.

Definition

- If A and B are square matrices, it is said that B **is similar to** A if there is an invertible matrix P such that $B = P^{-1}AP$.

Laplace Transform

Introduction:

The main objective of this module is to learn new methods to solve differential equations, in particular initial value problems. Essentially this is a technique which converts differential equations to algebraic equations which are easier to solve and then interpret this solution as the solution of the original differential equation.

Definition:

Laplace transform of a real function f(t) is defined as $L\{F(t)\} = \int_0^{\infty} e^{-st} F(t)dt.$

The resulting integral is a function of the variable s. To emphasize this very often we use the notation $L\{F(t)\} = f(s).$ Since *st* occur as an exponent in the definition of the Laplace transform and *t* stands for time, we say that *s* is frequency. This is because physical quantities in exponent should have to be dimensionless. In general when a function is multiplied by a standard function and the product integrated over certain limits, one gets what is known as integral transforms. Laplace transform is an example of this when the standard function is exponential function. Fourier transform is another example you will study in this course. Integral transforms, in general, have nice properties which are exploited to solve differential equations.

Basic property of Laplace transforms

1. One of the most important properties of Laplace transform is that it is a linear transformation which means for two functions *F* and *G* and constants *a* and *b*

$$L\{aF(t) \pm bG(t)\} = aL\{F(t)\} \pm bL\{G(t)\}.$$

Some Standard results

$F(t)$	$L\{F(t)\} = f(s).$
1	$\frac{1}{s}$
e^{at}	$\frac{1}{s-a}$
$sinat$	$\frac{a}{s^2+a^2}$
$cosat$	$\frac{s}{s^2+a^2}$
$sinhat$	$\frac{a}{s^2-a^2}$
$coshat$	$\frac{s}{s^2-a^2}$
t^n	$\frac{n!}{s^{n+1}}$ **where n is positive integer**
t^n	$\frac{\gamma(n+1)}{s^{n+1}}$

Problems

1. Find the laplace transform of $9+3t^2-5e^{2t}+5e^{-10t}$

Sol^n: $L\{9+3t^2-5e^{2t}+2e^{-10t}\} = L\{1\}+3L\{t^2\}-5L\{e^{2t}\}+2L\{e^{-10t}\}$

$$=\frac{9}{s}+3\left(\frac{2}{s^3}\right)-5\left(\frac{1}{s-2}\right)+5\left(\frac{1}{s+10}\right).$$

2. Find the laplace transform of $5^t-3cos2t+4sinh3t$.

Sol^n: $L\{\ 5^t-3cos2t+4sinh3t\} = L\{5^t\}-3L\{cos2t\}+4L\{sinh3t\}$

$$=\frac{1}{s-log5}+3\left(\frac{s}{s^2+4}\right)-4\left(\frac{3}{s^2-9}\right).$$

3. Find the laplace transform of $cosh^2 3t$

Sol^n:

$$L\{cosh^2 3t\} = L\left\{\left(\frac{e^{3t}+e^{-3t}}{2}\right)^2\right\}$$
$$=L\left\{\frac{e^{6t}+e^{-6t}+2}{4}\right\}$$
$$=\frac{1}{4}\left\{\frac{1}{s-6}+\frac{1}{s+6}+\frac{2}{s}\right\}$$

4. Find the laplace transform of $cos4t\ cos2t$

Sol^n:

$$L\{cos4t\ cos2t\} = L\left\{\frac{1}{2}(\cos(6t)+\cos(2t))\right\}$$
$$=\frac{1}{2}L\{(\cos(6t)+\cos(2t)\}$$
$$=\frac{1}{4}\left\{\frac{s}{s^2+36}+\frac{s}{s^2+4}\right\}$$

5. Find the laplace transform of $sin6t\ cos2t$

Sol^n:

$$L\{sin6t\ cos2t\} = L\left\{\frac{1}{2}(\sin(8t)+\sin(4t))\right\}$$
$$=\frac{1}{2}L\{(\sin(8t)+\sin(4t))\}$$
$$=\frac{1}{4}\left\{\frac{8}{s^2+64}+\frac{4}{s^2+16}\right\}$$

6. Find the laplace transform of $cos^2(5t)$

Sol^n:

$$L\{cos^2(5t)\} = L\left\{\frac{1}{2}(1+cos10t)\right\}$$

$$= \frac{1}{2} L\{(1 + cos10t)\}$$
$$= \frac{1}{4}\left\{\frac{1}{s} + \frac{s}{s^2 + 100}\right\}$$

7. Find the laplace transform of $sin^3(3t)$

Sol^n: $$L\{cos^2(5t)\} = L\left\{\frac{1}{4}(3sint - sin3t)\right\}$$
$$= \frac{1}{4} L\{(3sint - sin3t)\}$$
$$= \frac{1}{4}\left\{\frac{3}{s^2 + 1} - \frac{3}{s^2 + 9}\right\}$$

8. Find the laplace transform of $sin^3(3t)$

Sol^n: $$L\{cos^2(5t)\} = L\left\{\frac{1}{4}(3sint - sin3t)\right\}$$
$$= \frac{1}{4} L\{(3sint - sin3t)\}$$
$$= \frac{1}{4}\left\{\frac{3}{s^2 + 1} - \frac{3}{s^2 + 9}\right\}$$

9. Find the laplace transform of $cost\ cos2t\ cos3t$

Sol^n: $$L\{cost\ cos2t\ cos3t\} = L\left\{\frac{1}{2}[\mathrm{cos3t} + \mathrm{cost}]\mathrm{cos3t}\right\}$$
$$= \frac{1}{2} L\{\mathrm{cos^2 3t} + \mathrm{cos3t\ cost}\}$$
$$= \frac{1}{2} L\left\{\frac{1 + cos6t}{2} + \frac{1}{2}[\mathrm{cos4t} + \mathrm{cos2t}]\right\}$$
$$= \frac{1}{4} L\{1 + cos6t + cos4t + cos2t\}$$

$$= \frac{1}{4}\left[\frac{1}{s} + \frac{s}{s^2 + 36} + \frac{s}{s^2 + 16} + \frac{s}{s^2 + 4}\right]$$

10. Find the laplace transform of $sint\ sin2t\ sin3t$

Sol^n: $$L\{sint\ sin2t\ sin3t\} = L\left\{\frac{1}{2}[\mathrm{cost} - \mathrm{cos3t}]\mathrm{sin3t}\right\}$$
$$= \frac{1}{2} L\{\mathrm{cost\ sin3t} - \mathrm{cos3t\ sin3t}\}$$
$$= \frac{1}{2} L\left\{\frac{1}{2}[\mathrm{sin4t} + \mathrm{sin2t}] - \frac{sin6t}{2}\right\}$$
$$= \frac{1}{4} L\{sin4t + sin2t - sin6t\}$$

$$=\frac{1}{4}\left[\frac{4}{s^2+16}+\frac{2}{s^2+4}-\frac{6}{s^2+36}\right]$$

Properties of Laplace transform

1. Shifting property: If $L\{F(t)\}=f(s)$ then $L\{e^{at}F(t)\}=f(s-a)$.
2. If $L\{F(t)\}=f(s)$ then $L\{t^nF(t)\}=(-1)^n f^n(s)$.
3. If $L\{F(t)\}=f(s)$ then $L\left\{\frac{F(t)}{t}\right\}=\int_s^\infty f(s)\,ds$.

Problems

Find the Laplace transform of the following functions

1. $L\{e^{4t}cos2t\}$

Sol^n: $w.k.t\ \ L\{cos2t)=\frac{s}{s^2+4}$

$$L\{e^{4t}cos2t)=\frac{s-4}{(s-4)^2+4}\ .$$

2. $L\{e^{-2t}cos^2(5t)\}$

Sol^n: $w.k.t\ :\ L\{cos^2(5t)\}=L\left\{\frac{1}{2}(1+cos10t)\right\}$

$$=\frac{1}{2}L\{(1+cos10t)\}$$

$$=\frac{1}{4}\left\{\frac{1}{s}+\frac{s}{s^2+100}\right\}$$

$$L\{e^{-2t}cos^2(5t)\} = \frac{1}{4}\left\{\frac{1}{s+2} + \frac{s+2}{(s+2)^2+100}\right\}$$

3. $L\{e^{9t}sin6t\,cos2t\}$

Sol^n: $w.k.t$: $L\{sin6t\,cos2t\} = L\left\{\frac{1}{2}(\sin(8t)+\sin(4t))\right\}$

$$= \frac{1}{2}L\{(\sin(8t)+\sin(4t))\}$$

$$= \frac{1}{4}\left\{\frac{8}{s^2+64} + \frac{4}{s^2+16}\right\}$$

$$L\{e^{9t}sin6t\,cos2t\} = \frac{1}{4}\left\{\frac{8}{(s-9)^2+64} + \frac{4}{(s-9)^2+16}\right\}$$

4. $L\{sinhat\ cosat\}$

Sol^n: $L\{sinhat\ cosat\} = L\left\{\left(\frac{e^{at}-e^{-at}}{2}\right)\ cosat\right\}$

$$= L\left\{\left(\frac{e^{at}cosat - e^{-at}cosat}{2}\right)\right\}$$

$$= \frac{1}{2}L\{e^{at}cosat - e^{-at}cosat\}$$

$$= \frac{1}{2}\left\{\frac{s-a}{(s-a)^2+a^2} + \frac{s+a}{(s+a)^2+a^2}\right\}$$

5. $L\{tcos2t\}$

Sol^n: $w.k.t$ $L\{cos2t) = \frac{\mathrm{s}}{\mathrm{s}^2+4}$

$$L\{tcos2t) = (-1)\left(\frac{(\mathrm{s}^2+4)(1) - \mathrm{s}(2\mathrm{s})}{(\mathrm{s}^2+4)^2}\right)$$

$$L\{tcos2t) = (-1)\left(\frac{-\mathrm{s}^2+4}{(\mathrm{s}^2+4)^2}\right)$$

$$L\{tcos2t) = \left(\frac{\mathrm{s}^2-4}{(\mathrm{s}^2+4)^2}\right)$$

6. $L\{te^{-2t}sin4t\}$

Sol^n: $w.k.t$ $L\{sin4t) = \frac{4}{\mathrm{s}^2+16}$

$$L\{tsin4t) = (-1)\left(\frac{-4(2\mathrm{s})}{(\mathrm{s}^2+16)^2}\right)$$

$$L\{tsin4t) = (-1)\left(\frac{-8\mathrm{s}}{(\mathrm{s}^2+4)^2}\right)$$

$$L\{te^{-2t}sin4t\} = \left(\frac{8(\mathrm{s}+2)}{((\mathrm{s}+2)^2+4)^2}\right)$$

7. $L\{tsin6t\ cos2t\}$

Sol$^{\mathrm{n}}$: $w.k.t$: $L\{sin6t\ cos2t\} = L\left\{\frac{1}{2}(\sin(8t)+\sin(4t))\right\}$

$$=\frac{1}{2}L\{(\sin(8t)+\sin(4t))\}$$

$$=\frac{1}{4}\left\{\frac{8}{s^2+64}+\frac{4}{s^2+16}\right\}$$

$$L\{tsin6t\ cos2t\} = -\frac{1}{2}\left\{\frac{7s}{(\mathrm{s}^2+49)^2}+\frac{5}{(\mathrm{s}^2+25)^2}\right\}$$

8. $L\left\{\frac{e^{-at}+e^{-bt}}{t}\right\}$

Sol$^{\mathrm{n}}$: $w.k.t$: $L\{e^{-at}+e^{-bt}\}= \frac{7s}{(\mathrm{s}^2+49)^2}+\frac{5}{(\mathrm{s}^2+25)^2}$

$$L\left\{\frac{e^{-at}+e^{-bt}}{t}\right\}= \int_s^\infty \frac{7s}{(\mathrm{s}^2+49)^2}+\frac{5}{(\mathrm{s}^2+25)^2}\,ds$$

$$= [\log(s+a)-\log(s+b)]_s{}^{\infty}$$

$$=\log\left(\frac{s+b}{s+a}\right)$$

9. $L\left\{\frac{cosat+cosbt}{t}\right\}$

Sol$^{\mathrm{n}}$: $w.k.t$: $L\{cosat+cosbt\}= \frac{s}{s^2+a^2}+\frac{s}{s^2+b^2}$

$$L\left\{\frac{e^{-at}+e^{-bt}}{t}\right\}= \int_s^\infty \frac{s}{s^2+a^2}+\frac{s}{s^2+b^2}\,ds$$

$$=\frac{1}{2}\int_s^\infty \frac{2s}{s^2+a^2}+\frac{2s}{s^2+b^2}\,ds$$

$$=\frac{1}{2}\log\left(\frac{s^2+a^2}{s^2+b^2}\right)_s^{\infty}$$

$$=\frac{1}{2}\log\left(\frac{s^2+b^2}{s^2+a^2}\right)$$

10. $L\left\{\frac{sinkt}{t}\right\}$

Sol^n: $w.k.t$: $L\{sinkt\}=\frac{k}{s^2+k^2}$

$$L\left\{\frac{sinkt}{t}\right\}=\int_s^{\infty}\frac{k}{s^2+k^2}ds$$

$$=\text{k}\frac{1}{k}[tan^{-1}s]_s^{\infty}$$

$$=(tan^{-1}\infty-tan^{-1}s)$$

$$=(\frac{\pi}{2}-tan^{-1}s)$$

$$=cot^{-1}s$$

Periodic function

A function F(t) is said to be a periodic function of period "T" if F(t+T)=F(t)
Eg(i): "sin t" is a periodic function of period "2 π"

Theorem

Statement: If $F(t)$ is a periodic function of period T then, is defined as

$$L\{F(t)\}=\frac{1}{1-e^{-ST}}\int_0^T e^{-st}F(t)dt.$$

Problems

1. A periodic function of period $\frac{2\pi}{w}$ is defined by

$$F(t)=\begin{cases} Esinwt & 0\le t\le\frac{\pi}{w} \\ 0 & \frac{\pi}{w}\le t\le\frac{2\pi}{w}\end{cases}$$

Where E & w are constants then show that

$$L\{F(t)\} = \frac{Ew}{(s^2+w^2)(1-e^{-\frac{\pi s}{w}})}$$

Sol^{n}: $w.k.t \quad L\{F(t)\} = \frac{1}{1-e^{-ST}}\int_0^{\infty} e^{-st}\, F(t)dt.,$ $\qquad T = \frac{2\pi}{w}$

$$= \frac{1}{1-e^{-\frac{2\pi S}{w}}}\int_0^{\frac{2\pi}{w}} e^{-st}\, F(t)dt$$

$$= \frac{1}{1-e^{-\frac{2\pi S}{w}}}\left\{\int_0^{\frac{\pi}{w}} e^{-st}\, F(t)dt + \int_{\frac{\pi}{w}}^{\frac{2\pi}{w}} e^{-st}\, F(t)dt\right\}$$

$$= \frac{1}{1-e^{-\frac{2\pi S}{w}}}\left\{\int_0^{\frac{\pi}{w}} e^{-st}\, Esinwtdt + \int_{\frac{\pi}{w}}^{\frac{2\pi}{w}} e^{-st}\, 0dt\right\}$$

$$= \frac{1}{1-e^{-\frac{2\pi S}{w}}}\left\{E\int_0^{\frac{\pi}{w}} e^{-st}\, sinwtdt\right\}$$

$$= \frac{E}{1-e^{-\frac{2\pi S}{w}}}\left\{\frac{e^{-st}}{s^2+w^2}(-s\,sinwt - wcoswt)\right\}_0^{\frac{\pi}{w}}$$

$$= \frac{Ew}{\left(1-e^{-\frac{2\pi S}{w}}\right)(s^2+w^2)}\left[1+e^{-\frac{\pi s}{w}}\right]$$

$$= \frac{Ew}{\left(1-\left(e^{-\frac{\pi S}{w}}\right)^2\right)(s^2+w^2)}\left[1+e^{-\frac{\pi s}{w}}\right]$$

$$L\{F(t)\} = \frac{Ew}{(s^2 + w^2)(1 - e^{-\frac{\pi s}{w}})}$$

2. A periodic function of period $2a$ is defined by $F(t) =$
$$\begin{cases} t & 0 \le t \le a \\ 2a - t & a \le t \le 2a \end{cases}$$

Then prove that $L\{F(t)\} = \frac{tanh\left(\frac{as}{2}\right)}{s^2}$

Sol^n: $w.k.t \quad L\{F(t)\} = \frac{1}{1-e^{-ST}} \int_0^T e^{-st} F(t)dt.,$ $\qquad T = 2a$

$$= \frac{1}{1 - e^{-2as}} \int_0^T e^{-st} F(t)dt$$

$$= \frac{1}{1 - e^{-2as}} \left\{ \int_0^a e^{-st} F(t)dt + \int_a^{2a} e^{-st} F(t)dt \right\}$$

$$= \frac{1}{1-e^{-2as}} \left\{ \int_0^a e^{-st}\, t\, dt + \int_a^{2a} e^{-st} (2a - t)dt \right\}$$

$$= \frac{1}{1-e^{-2as}} \left(\left[\frac{e^{-st}}{-s}\right]_0^a + \left[e^{-st}(-1) - (2a - t)\frac{e^{-st}}{-s} \right]_a^{2a} \right)$$

$$= \frac{1}{1 - e^{-2as}} \left(\frac{1 + (e^{-as})^2 - 2e^{-as}}{s^2} \right)$$

$$= \frac{1}{s^2} \left[\frac{1 - e^{-as}}{1 + e^{-as}} \right]$$

$$L\{F(t)\} = \frac{tanh\left(\frac{as}{2}\right)}{s^2}$$

Exercise

1. A periodic function of period a is defined by $F(t) = \begin{cases} E & 0 \le t \le \frac{a}{2} \\ -E & \frac{a}{2} \le t \le a \end{cases}$ then find $L\{F(t)\}$.
2. Find the laplace ransform of $F(t) = E\sin\left(\frac{\pi t}{w}\right)$ $0 \le t \le w$ given that $F(t+w) = F(t)$.

Unit step function (or) Heaviside functions

The unit step function or Heaviside function $U(t-a)$ is defined as

$$U(t-a) = \begin{cases} 0 & t \le a \\ 1 & t > a \end{cases}.$$

In particular, when a=0, $U(t-a) = \begin{cases} 0 & t \le 0 \\ 1 & t > 0 \end{cases}$

Properties

1. $L\{U(t-a)\} = \frac{e^{-as}}{s}$
2. If $L\{F(t)\} = f(s)$ then $L\{F(t-a).U(t-a)\} = e^{-as}f(s) = e^{-as}L\{F(t)\}$

Problems

I Find the Laplace transform of the following functions

1. $[e^{t-1} + \sin(t-1)]U(t-1)$

Sol^n: Let $F(t-1) = [e^{t-1} + \sin(t-1)]U(t-1)$

Replace t by t+1

$F(t) = [e^t + \sin(\mathrm{t})]\mathrm{U(t)}$

$L\{F(t)\} = L\{[e^t + \sin(\mathrm{t})]\mathrm{U(t)}\}$

$$L\{F(t)\} = \frac{1}{s-1} + \frac{1}{s^2+1}$$

$$L\{F(t-1)\} = e^{-s}\, L\{F(t)\}$$

$$L\{F(t-1)\} = e^{-s}\left(\frac{1}{s-1} + \frac{1}{s^2+1}\right)$$

2. $[e^{-2t}\mathrm{U(t-1)}]$

Solⁿ: Let $F(t-1) = [e^{-2t}]$

Replace t by t+1

$$F(t) = \left[e^{-2t(t+1)}\right]$$

$$L\{F(t)\} = e^{-2}\,\tfrac{1}{s+2}$$

$$L\{F(t-1)U(t-1)\} = e^{-s}\, L\{F(t)\}$$

$$L\{e^{-2t}U(t-1)\} = e^{-s}\left(\frac{e^{-2}}{s+2}\right)$$

Note 1:If $F(t) = \begin{cases} f_1(t) & 0 < t < a \\ f_2(t) & t > a \end{cases}$

then $F(t) = f_1(t) + [f_2(t) - f_1(t)]U(t-a)$

Note 2:If $F(t) = \begin{cases} f_1(t) & 0 < t < a \\ f_2(t) & a < t < b \\ f_3(t) & t > b \end{cases}$

then $F(t) = f_1(t) + [f_2(t) - f_1(t)]U(t-a) + [f_3(t) - f_2(t)]U(t-b)$

Express the function in terms of unit step function and hence find its Laplace Transform

1. F(t) = $\begin{cases} t & 0 < t < 2 \\ t^2 & t > 2 \end{cases}$

Soln. F(t)= t+(t^2-t). $U(t-2)$

L{F(t)}= L{t}+ L{(t^2 -t) . $U(t-2)$}

Consider,

L{(t^2 -t) . $U(t-2)$}

Let, F(t-2) = t^2 -t

F(t)= $(t+2)^2$ - (t+2)

F(t)= t^2 +4+4t-t-2

F(t)=t^2 +3t+2

L{F(t)} = $\frac{2}{s^3} + \frac{3}{s^2} + \frac{2}{s}$

We know that,

L{F(t-a) . $U(t-a)$} = e^{-as}.L{F(t)}

Put a=2,

L{F(t-2). . $U(t-2)$} = e^{-2s}.L{F(t)}

L{F(t^2-t). . $U(t-2)$} =$e^{-2s} . \frac{2}{s^3} + \frac{3}{s^2} + \frac{2}{s}$

∴ L{F(t)} =$\frac{1}{s^2} + e^{-2s} . \frac{2}{s^3} + \frac{3}{s^2} + \frac{2}{s}$

2. Express F (t) = $\begin{cases} cos\ t & 0 < t < \pi \\ cos2t & \pi < \ t > 2\pi \\ cos\ 3t & t > 2\pi \end{cases}$

Terms of Heaviside function and hence find L {F (t)}.

Soln.

F(t) = cos t + (cos 2t – cos t). ∪ $(t-\pi)+(\cos 3t-\cos 2t)$.∪ $(t-2\pi)$

L{F(t)} = L{cos t} + L {(cos 2t – cos t) . ∪ $(t-\pi)+L\,\{(\cos 3t-\cos 2t)$.∪ $(t-2\pi)\}$

L{F(t)} = $\frac{s}{(\mathbf{s2}+1)}$+ L {(cos 2t – cos t) . ∪ $(t-\pi)+L\,\{(\cos 3t-\cos 2t)$.∪ $(t-2\pi)\}$

Consider,

$$L\,\{(\cos 2t-\cos t).\cup\,(t-\pi)\}$$

Let,

$F(t-\pi)$ = cos 2t – cos t

F(t) = cos 2(t – π) - cos (t – π)

F(t) = cos (2t – 2π) - cos (π – t)

F(t) = cos 2t + cos t

L{(t)} = $\frac{s}{(s^2+4)}+\frac{s}{(s^2+1)}$

We know that,

L{F(t-a) . $U(t-a)\}$ = e^{-as}.L{F(t)}

Put a= π,

L{F(t – π**)** . $U(t-\pi)\}$ = $e^{-\pi s}$. L{F(t)}

$L\{(\cos 2t - \cos t).\cup(t-\pi)\} = e^{-\pi s}.\left[\frac{s}{(s^2+4)} + \frac{s}{(s^2+1)}\right]$

Consider,

$L\{(\cos 3t - \cos 2t).\cup(t-2\pi)\}$

Let,

$F(t-2\pi)$ = cos 3t – cos 2t

F(t) = cos 3(t – 2π) – cos 2 (t – 2π)

F(t) = cos (6 π +2t) - cos (4π + $2t$)

F(t) = cos 3t - cos $2t$

L{(t)} = $\frac{s}{(s^2+9)} + \frac{s}{(s^2+4)}$

We know that,

L{F(t-a) . $U(t-a)$} = e^{-as}.L{F(s)}

Put a= 2π,

L{F(t – 2π) . $U(t-2\pi)$} = $e^{-2\pi s}$. L{F(t)}

$L\{(\cos 3t - \cos 2t).\cup(t-2\pi)\} = e^{-2\pi s}.\left[\frac{s}{(s^2+9)} + \frac{s}{(s^2+4)}\right]$

L{F(t)} = $\frac{s}{(\mathbf{s2}+1)} + e^{-2\pi s}\left[\frac{s}{(s^2+4)} + \frac{s}{(s^2+1)}\right] + e^{-2\pi s}\left[\frac{s}{(s^2+9)} + \frac{s}{(s^2+4)}\right]$

Inverse Laplace Transform

If L{F(t)} =f(s) then F(t) is called the inverse Laplace transform of f(s) and is denoted by $L^{-1}[f(s)]$.

Basic properties

1. $L^{-1}[f(s) \pm g(s)] = L^{-1}[f(s)] \pm L^{-1}[g(s)]$
2. $L^{-1}[kf(s)] = k\, L^{-1}[f(s)]$

Some Standard results

$f(s)$	$L^{-1}[f(s)] = F(t).$
$\frac{1}{s}$	1
$\frac{1}{s-a}$	e^{at}
$\frac{a}{s^2+a^2}$	$sinat$
$\frac{s}{s^2+a^2}$	$cosat$
$\frac{a}{s^2-a^2}$	$sinhat$
$\frac{s}{s^2-a^2}$	$coshat$
$\frac{n!}{s^{n+1}}$ where n is positive integer	t^n
$\frac{\gamma(n+1)}{s^{n+1}}$	t^n

Problems

Find the inverse Laplace transform of the following functions

1. $L^{-1}\left[\frac{1}{s+2}+\frac{3}{2s+5}-\frac{4}{3s-2}\right]$

$$= L^{-1}\left[\frac{1}{s+2}\right]+L^{-1}\left[\frac{3}{2s+5}\right]-L^{-1}\left[\frac{4}{3s-2}\right]$$

$$e^{-2t}+\frac{3}{2}e^{-\frac{5}{2}t}-\frac{4}{3}e^{\frac{2}{3}t}.$$

2. $L^{-1}\left[\frac{s+2}{s^2+36}+\frac{4s-1}{s^2+25}\right]$

$$= L^{-1}\left[\frac{s}{s^2+36}\right]+L^{-1}\left[\frac{2}{s^2+36}\right]+\ L^{-1}\left[\frac{4s}{s^2+25}\right]-L^{-1}\left[\frac{1}{s^2+25}\right]$$

$$= cos6t+\frac{sin6t}{3}+4cos5t-\frac{sin5t}{3}$$

3. $L^{-1}\left[\frac{(s+2)^3}{s^6}\right]$

$$= L^{-1}\left[\frac{s^3+8+6s^2+12s}{s^6}\right]$$

$$= L^{-1}\left[\frac{s^3}{s^6}+\frac{8}{s^6}+\frac{6s^2}{s^6}+\frac{12s}{s^6}\right]$$

$$= \frac{t^2}{2!}+8\frac{t^5}{5!}+6\frac{t^3}{3!}+12\frac{t^4}{4!}$$

Inverse Laplace transform of $\mathbf{f(s-a)}$

$w.k.t\ \ if\ \ L\{F(t)\} = f(s)$ then, $L\{e^{at}F(t)\} = f(s-a)$

$If\ \ L^{-1}[f(s] = F(t)$ then $L^{-1}\{f(s-a)\} = e^{at}L^{-1}[f(s]$

Problems

Find the inverse Laplace transform of the following functions

1. $L^{-1}\left[\frac{1}{(s-3)^4}\right]$

$= e^{3t}L^{-1}\left[\frac{1}{(s)^4}\right]$
$=e^{3t}\left[\frac{t^3}{3!}\right]$

2. $L^{-1}\left[\frac{s-3}{(s-3)^2+9}\right]$

$= e^{3t}L^{-1}\left[\frac{s}{s^2+9}\right]$
$=e^{3t}cos3t$

3. $L^{-1}\left[\frac{s-3}{(s-3)^2+9}\right]$

$= e^{3t}L^{-1}\left[\frac{s}{s^2+9}\right]$
$=e^{3t}cos3t$

4. $L^{-1}\left[\frac{s+1}{s^2+2s+5}\right]$

$= L^{-1}\left[\frac{s+1}{(s+1)^2+4}\right]$
$=e^{-t}\ L^{-1}\left[\frac{s}{s^2+4}\right]$
$=e^{-t}cos2t.$

5. $L^{-1}\left[\frac{1}{s^2+3s+2}\right]$

$= L^{-1}\left[\frac{1}{(s+2)(s+1)}\right]$
$= L^{-1}\left[\frac{-1}{(s+2)}+\frac{1}{(s+1)}\right]$
$=-e^{-2t}+-e^{-t}.$

6. $L^{-1}\left[\frac{1}{(s+2)(s+1)(s+3)}\right]$

$$= L^{-1}\left[\frac{\frac{1}{2}}{(s+1)} + \frac{-1}{(s+2)} + \frac{\frac{1}{2}}{(s+3)}\right]$$
$$=\frac{1}{2}\, L^{-1}\left[\frac{1}{(s+1)}\right] - \, L^{-1}\left[\frac{1}{(s+2)}\right] + \frac{1}{2}\, L^{-1}\left[\frac{1}{(s+3)}\right]$$
$$=\frac{1}{2}e^{-t} - e^{-2t} + \frac{1}{2}e^{-3t}.$$

7. $L^{-1}\left[\frac{4s+5}{(s+1)^2(s+2)}\right]$

$$= L^{-1}\left[\frac{1}{(s+1)^2} + \frac{3}{(s+1)} + \frac{-3}{(s+2)}\right]$$
$$= L^{-1}\left[\frac{1}{(s+1)^2}\right] - \, L^{-1}\left[\frac{3}{(s+1)}\right] + \, L^{-1}\left[\frac{-3}{(s+2)}\right]$$
$$=te^{-t}+3e^{-t} - e^{-2t}.$$

8. $L^{-1}\left[\frac{2s^2+5s-4}{s^3+s^2-2s}\right]$

$$= L^{-1}\left[\frac{2s^2+5s-4}{s(s^2+s-2)}\right]$$
$$= L^{-1}\left[\frac{2}{s} + \frac{-1}{(s+2)} + \frac{1}{(s-1)}\right]$$
$$=2 - e^{-2t} + e^{t}.$$

Inverse Laplace transform of logarithmic functions

Find the inverse Laplace transform of the following functions

1. $L^{-1}\left[log\left(\frac{s+a}{s+b}\right)\right]$

Sol^{n}: Let $L^{-1}\left[log\left(\frac{s+a}{s+b}\right)\right]=F(t)$

$$L\{F(t)\} = log\left(\frac{s+a}{s+b}\right)$$
$$= \log(s+a) - \log(s+b)$$

$$L\{tF(t)\} = (-1)\left[\frac{1}{s+a} + \frac{1}{s+b}\right]$$

$$= \left[\frac{1}{s+b} - \frac{1}{s+a}\right]$$
$$tF(t) = L^{-1}\left[\frac{1}{s+b} - \frac{1}{s+a}\right]$$
$$tF(t) = e^{-bt} - e^{-at}$$
$$F(t) = \frac{e^{-bt} - e^{-at}}{t}$$
$$L^{-1}\left[log\left(\frac{s+a}{s+b}\right)\right] = \frac{e^{-bt}-e^{-at}}{t}.$$

2. $L^{-1}\left[log\left(\frac{s}{s-1}\right)\right]$

Sol^{n}: Let $L^{-1}\left[log\left(\frac{s}{s-1}\right)\right]$=$F(t)$

$$L\{F(t)\} = log\left(\frac{s}{s-1}\right)$$
$$= \log(s) - \log(s-1)$$

$$L\{tF(t)\} = (-1)\left[\frac{1}{s} + \frac{1}{s-1}\right]$$

$$= \left[\frac{1}{s-1} - \frac{1}{s}\right]$$
$$tF(t) = L^{-1}\left[\frac{1}{s-1} - \frac{1}{s}\right]$$
$$tF(t) = e^{t} - 1$$
$$F(t) = \frac{e^{t} - 1}{t}$$
$$L^{-1}\left[log\left(\frac{s}{s-1}\right)\right] = \frac{e^{t}-1}{t}.$$

3. $L^{-1}\left[log\left(\frac{1-s^2}{s^2}\right)\right]$

Sol^{n}: Let $L^{-1}\left[log\left(\frac{s}{s-1}\right)\right]=F(t)$

$$L\{F(t)\} = log\left(\frac{1-s^2}{s^2}\right)$$
$$= \log(1-s^2) - \log(s^2)$$

$$L\{tF(t)\} = (-1)\left[\frac{1}{1-s^2}(-2s) + \frac{1}{s^2}(2s)\right]$$

$$= \left[\frac{2s}{s^2} - \frac{2s}{s^2-1}\right]$$
$$tF(t) = 2L^{-1}\left[\frac{1}{s} - \frac{s}{s^2-1}\right]$$
$$tF(t) = 2[1 - cosht]$$

$$F(t) = \frac{2[1-cosht]}{t}$$
$$L^{-1}\left[log\left(\frac{\mathbf{1-s^2}}{\mathbf{s^2}}\right)\right] = \frac{2[1-cosht]}{t}.$$

Inverse Laplace transform of $e^{-as}f(s)$

$w.k.t \quad if \;\; L\{F(t)\} = f(s)$ then, $L\{F(t-a).U(t-a)\} = e^{-as}f(s)$

$If \;\; L^{-1}[f(s] = F(t)$ then, $L^{-1}\{e^{-as}f(s)\} = F(t-a).U(t-a)$

Find the inverse Laplace transform of the following functions

1. $L^{-1}\left[\frac{e^{-2s}}{s^2+1}\right]$

$$w.k.t \qquad L^{-1}\left[\frac{1}{s^2+1}\right] = sint$$

$$L^{-1}\left[\frac{e^{-2s}}{s^2+1}\right] = \sin(t-2)\,U(t-2)$$

2. $L^{-1}\left[\frac{se^{-3s}}{s^2+4}\right]$

$$w.k.t \qquad L^{-1}\left[\frac{s}{s^2+4}\right] = cos2t$$

$$L^{-1}\left[\frac{se^{-3s}}{s^2+4}\right] = cos2(-3)t\,U(t-3).$$

3. $L^{-1}\left[\frac{e^{-3s}}{(s+1)^4}\right]$

$$w.k.t \qquad L^{-1}\left[\frac{1}{(s+1)^4}\right] = e^{-t}\,\frac{t^3}{3!}$$

$$L^{-1}\left[\frac{e^{-3s}}{(s+1)^4}\right] = e^{-(t-3)}\,\frac{(t-3)^3}{3!}\,U(t-3)$$

4. $L^{-1}\left[\frac{e^{-4s}}{(s+1)(s^2+2s+2)}\right]$

$$w.k.t \qquad L^{-1}\left[\frac{1}{(s+1)(s^2+2s+2)}\right] = \; L^{-1}\left[\frac{1}{(s+1)((s+1)^2+1)}\right]$$

$$= e^{-t} \qquad L^{-1}\left[\frac{1}{(s)(s^2+1)}\right]$$

$$= e^{-t} \qquad L^{-1}\left[\frac{1}{(s)} + \frac{-s}{(s^2+1)}\right]$$

$$= e^{-t}\,[1 - cost]$$

$$L^{-1}\left[\frac{e^{-4s}}{(s+1)(s^2+2s+2)}\right] = e^{-(t-4)}\,[1-\cos(t-4)]U(t-4)$$

Convolution

The convolution of 2 functions F(t) & G(t) usually denoted by F(t).G(t). and as defined in the form of an integral as follows.

F(t).G(t) = $\int_{u=0}^{t} F(u).G(t-u)du$

Convolutions theorem

If L^{-1}[f(s).g(s)] = F(t) and L^{-1}[g(s)] = G(t) then,

L^{-1}[f(s).g(s)] = $\int_{u=0}^{t} F(u).G(t-u)du$

Problems:

Using convolution theorem obtains the inverse Laplace transform of the following functions.

1. $L^{-1}\left[\frac{1}{s(s^2+a)}\right]$

Sol^n: Let f(s) = $\frac{1}{s}$, g(s) = $\frac{1}{s^2+a^2}$

L^{-1}[f(s)] = 1 = F(t) , L^{-1}[g(s)] = $\frac{\sin at}{a}$ = G(t)

WKT,

$$L^{-1}[f(s).g(s)] = \int_{u=0}^{t} F(u).G(t-u)du$$

$$L^{-1}\left[\frac{1}{s}\,\frac{1}{s^2+a^2}\right] = \int_{u=0}^{t} 1\ \frac{\sin a(t-u)}{a}du$$

$$L^{-1}\left[\frac{1}{s(s^2+a^2)}\right] = \frac{1}{a}\int_{u=0}^{t} sin(at-au)du$$

$$= \frac{1}{a}\left[-\frac{\cos(at-au)}{-a}\right]_0^t$$

$$= \frac{1}{a^2}\left[\cos(at-au)\right]_0^t$$

$$= \frac{1}{a^2}[1-\cos at]$$

2. $L^{-1}\left[\frac{1}{(s^2+a^2)^2}\right]$

Sol^n: Let f(s) = $\frac{1}{s^2+a^2}$, g(s) = $\frac{1}{s^2+a^2}$

$$L^{-1}[f(s)] = \frac{\sin at}{a} = F(t) \;, \qquad L^{-1}[g(s)] = \frac{\sin at}{a} = G(t)$$

WKT,

$$L^{-1}[f(s).g(s)] = \int_{u=0}^{t} F(u).G(t-u)du$$

$$L^{-1}\left[\frac{1}{s^2+a^2}\,\frac{1}{s^2+a^2}\right] = \int_{u=0}^{t} \frac{\sin au}{a}\,\frac{\sin a(t-u)}{a}du$$

$$L^{-1}\left[\frac{1}{s(s^2+a^2)}\right] = \frac{1}{a^2}\int_{u=0}^{t} \sin au\ \ sin(at-au)du$$

$$= -\frac{1}{2a^2}\int_{u=0}^{t} \cos at - \ cos(2au-at)du$$

$$= -\frac{1}{2a^2}\left[-\text{u}\cos at - \frac{sin(2au-at)}{2a}\right]_0^t$$

$$= \frac{1}{2a^3}[\text{sinat} - \text{at cosat}]$$

Laplace transform of derivative

Note *(i)* $L\{y^1(t)\} = s\ L\{y(t)\} - y(0)$
(ii) $L\{y^{11}(t)\} = s^2\ L\{y(t)\} - sy(0) - y^1(0)$
(Iii) $L\{y^{111}(t)\} = s^3\ L\{y(t)\} - s^2\ y(0) - sy^1(0) - y^{11}(0)$

Problems

1. Solve by using Laplace transform method $\frac{d^2y}{dt^2} + k^2y = 0$ given that $y(0) = 2,\ y^1(0) = 0.$

Sol^n: Given

$$y^{11}(t)+k^2y=0$$
$$L\{y^{11}(t)\}+k^2L\{y(t)\}=0$$
$$s^2L\{y(t)\}-s\,y(0)-y^1(0)+k^2L\{y(t)\}=0$$
$$s^2L\{y(t)\}-2s\ -0+k^2L\{y(t)\}=0$$
$$(s^2+k^2)\ L\{y(t)\}=2s$$
$$L\{y(t)\}=\frac{2s}{(s^2+k^2)}$$
$$y(t)=\ L^{-1}\left[\frac{2s}{(s^2+k^2)}\right]$$
$$y(t)=2coskt$$

2. Solve by using Laplace transform method

$$y^{111}(t)+2y^{11}(t)-2y^1(t)+2y(t)=0$$
given that $y(0)=y^1(0)=0,\ y^{11}(0)=6$

Sol^{n}: Given

$$y^{111}(t)+2y^{11}(t)-2y^1(t)+2y(t)=0$$
$$L\{y^{111}(t)\}+2L\{y^{11}(t)\}-2L\{y^1(t)\}+2L\{y(t)=0$$
$$s^3L\{y(t)\}-s^2\,y(0)-sy^1(0)-y^{11}(0)+2[s^2L\{y(t)\}-s\,y(0)-$$
$$y^1(0)]-2[sL\{y(t)\}-y(0)]+2L\{y(t)\}=0$$
$$(s^3+2s^2-s-2)L\{y(t)\}-6=0$$
$$(s^3+2s^2-s-2)L\{y(t)\}=6$$
$$L\{y(t)\}=\frac{6}{(s^3+2s^2-s-2)}$$
$$y(t)=\ L^{-1}\left[\frac{6}{(s^3+2s^2-s-2)}\right]$$
$$y(t)=\ L^{-1}\left[\frac{6}{(s^3+2s^2-s-2)}\right]$$
$$y(t)=\ L^{-1}\left[\frac{-3}{(s+1)}+\frac{1}{(s-1)}+\frac{2}{(s+2)}\right]$$
$$y(t)=-3e^{-t}+e^{t}+2e^{-2t}$$

NUMERICAL METHODS-I

SOLUTION 0F ALGEBRAIC & TRANSCENDENTAL EQUATIONS

An equation $f(x) = 0$ is classified into 2 types

(1) Algebraic or Polynomial

(2) Transcendental

$f(x) = a_0x^n + a_1x^{n-1} + \cdots\cdots\cdots\cdots + a_{n-1}x + a_n,\ a_0 \neq 0$ is called algebraic or polynomial function of degree n. A function which is sum, difference or product of 2 polynomials is called an algebraic function, otherwise, the function is called a transcendental function. For example $e^x,\ \log x,\ \cos x$ etc are transcendental functions.

[*i.e.* $f(x) = \sum_{i=0}^{\infty} a_i x^i$ is called a transcendental equation.

Eg. $\cosh x + \sin x,\ e^x + x,\ 10 + \log x$ etc.]

Root of an equation: A number m (real or complex) is called a root (or a **solution**) of an equation $f(x) = 0$ if $f(m) = 0$. Then we say that m satisfies the equation $f(x) = 0$ or that m is a **zero** of the function $f(x)$. The process of finding roots of an equation $f(x) = 0$ is known as solving the equation $f(x) = 0$.

Note: The repeated execution of the same process where at each step the result of the preceding step is used is known as **iteration process**.

To find the roots of an equation $f(x) = 0$ we start with a known approximate solution and apply any of the following methods:

Regula – Falsi Method (Or Method of false position) :-

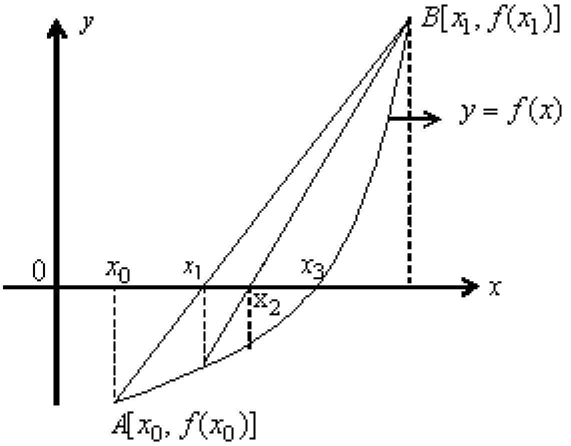
Here we choose two points x_0 & x_1 such that $f(x_0)$ & $f(x_1)$ are of opposite signs. *i.e.* the graph of $y = f(x)$ crosses the x – axis between these points.

Therefore a root lies between x_0 & x_1.

Now equation of the line joining the points $A[x_0, f(x_0)]$ & $B[x_1, f(x_1)]$ is

$$\frac{y - f(x_0)}{f(x_1) - f(x_0)} = \frac{x - x_0}{x_1 - x_0} \text{----- (1)}$$

In this method we replace the curve AB by the line AB and take the point of intersection of the line (i.e. eqn. (1)) with the x – axis as an approximation to the root.



$\therefore$ putting $y = 0$ in (1), we get

$$\frac{-f(x_0)}{f(x_1)-f(x_0)} = \frac{x-x_0}{x_1-x_0} \Rightarrow x = x_0 - \frac{x_1-x_0}{f(x_1)-f(x_0)} f(x_0)$$

i.e. $x_2 = x_0 - \dfrac{x_1-x_0}{f(x_1)-f(x_0)} f(x_0)$----- (2) which is an approximation to the root.

If now $f(x_0)$ & $f(x_2)$ are of opposite signs, then the root lies between x_0 & x_2. So replacing x_1 by x_2 in (2) we get the next approximation x_3. If $f(x_2)$ & $f(x_1)$ are of opposite signs, then the root lies between x_1 & x_2. Then get x_3 by replacing x_0 by x_2 in (2). Repeat the above process till the root is found to the desired accuracy.

The iteration process based on (1) is known as the **method of false position**.

Examples: **1)** Find the root of the equation $x^3 - 2x - 5 = 0$ by the method of false position correct to 3 decimal places.

Solution: Let $f(x) = x^3 - 2x - 5$, then

$f(2) = 2^3 - 2\times 2 - 5 = -1 < 0$ & $f(3) = 3^3 - 2\times 3 - 5 = 16 > 0$

$\therefore$ A root lies between 2 & 3. Let $x_0 = 2$ & $x_1 = 3$, then

$$x_2 = x_0 - \frac{x_1-x_0}{f(x_1)-f(x_0)} f(x_0) = 2 - \frac{3-2}{16-(-1)} \times (-1) = 2 + \frac{1}{17} = 2.0588$$

Now $f(x_2) = (2.0588)^3 - 2(2.0588) - 5 = -0.3911 < 0$

$\therefore$ The root lies between 2.0588 & 3. $\therefore$ We get

$$x_3 = x_2 - \frac{x_1-x_2}{f(x_1)-f(x_2)} f(x_2) = 2.0588 - \frac{3-2.0588}{16+0.3911} \times (-0.3912)$$

$$= 2.0588 + \frac{0.9412}{16.3911} \times 0.3911 = 2.0813$$

Repeating this process the successive approximations are $x_4 = 2.0896$, $x_5 = 2.0927$, $x_6 = 2.0938$, $x_7 = 2.0939$, $x_8 = 2.0943$ etc.

Hence the root is 2.094 correct to 3 decimal places.

---

2) Find areal root of the equation $x \log_{10} x = 1.2$ by regula - falsi method correct to 4 decimal places. **(VTU 2004)**

Let $f(x) = x \log_{10} x - 1.2$, then

$f(1) = 1 \times \log_{10} 1 - 1.2 = -1.2 < 0,\ f(2) = 2 \times \log_{10} 2 - 1.2 = -0.59794 < 0$

$f(3) = 3\log_{10} 3 - 1.2 = 0.23136 > 0$

$\therefore$ A root lies between 2 & 3.Taking $x_0 = 2 \,\&\, x_1 = 3$ in the regula - falsi method we get,

$$x_2 = x_0 - \frac{x_1 - x_0}{f(x_1) - f(x_0)} f(x_0) = 2 - \frac{3-2}{0.23136 - (-0.59794)} \times (-0.59794) = 2 + \frac{1}{0.8293} \times 0.59794$$

$$= 2.72102$$

Now $f(x_2) = 2.72102 \times \log_{10} 2.72102 - 1.2 = -0.01709 < 0$

i.e. root lies between 2.72102 & 3 .

$$\therefore x_3 = 2.72102 - \frac{3 - 2.72102}{0.23136 - (-0.01709)} \times (-0.01709) = 2.72102 + \frac{0.27898}{0.24845} \times 0.01709 = 2.74021$$

Repeating this process the successive approximations are $x_4 = 2.74024,\ x_5 = 2.74063$,

$x_6 = 2.74064$.

Hence the root is 2.7406 correct to 4 decimal places.

---

3) Find the real root of the equation $\cos x = 3x - 1$, correct to 3 decimal places by the method of False position. **[VTU 2003]**

Here $f(x) = \cos x - 3x + 1$

$\therefore\ f(0) = 2 > 0,\ f(1) = -1.4597 < 0 \Rightarrow$ Root lies between 0 & 1

Let $x_0 = 0 \,\&\, x_1 = 1$

$$x_2 = x_0 - \frac{x_1 - x_0}{f(x_1) - f(x_0)} f(x_0) = 0 - \frac{1-0}{-1.4597 - 2} \times 2 = 0 + \frac{2}{3.4597} = 0.5781$$

$f(x_2) = f(0.5781) = 0.1032 > 0$

Hence root lies between 0.5781 & 1

$$\therefore x_3 = 0.5781 - \frac{1 - 0.5781}{-1.4597 - 0.1032} \times 0.1032 = 0.5781 + \frac{0.4219}{1.5629} \times 0.1032 = 0.6060$$

$f(x_3) = 0.0039 > 0$

$\therefore$ root lies between 0.6060 & 1

$$x_4 = 0.6060 - \frac{1 - 0.6060}{-1.4597 - 0.0039} \times 0.0039 = 0.6060 + \frac{0.394}{1.4636} \times 0.0039 = 0.607$$

$f(x_4) = 0.000363 > 0$

Hence root lies between 0.607 & 1

$$x_5 = 0.607 - \frac{1-0.607}{-1.4597-0.000363}\times 0.000363 = 0.607 + \frac{0.393}{1.460063}\times 0.000363 = 0.6071$$

Hence the root is 0.607 correct to 3 decimal places.

---

4) Find the fourth root of 12 correct to 3 decimal places using the method of false position.

Answer: Let $x = (12)^{1/4} \Rightarrow x^4 - 12 = 0$. Take $f(x) = x^4 - 12$, then

$f(1) = 1-12 = -11 < 0,\ f(2) = 2^4 - 12 = 16-12 = 4 > 0$

$\therefore$ root lies between 1 & 2. Taking $x_0 = 1$ & $x_1 = 2$, we get $f(x_0) = -11$ & $f(x_1) = 4$

$$\therefore x_2 = x_0 - \frac{x_1 - x_0}{f(x_1) - f(x_0)} f(x_0) = 1 - \frac{2-1}{4-(-11)}\times(-11) = 1 + \frac{1}{15}\times 11 = 1.7333$$

Now $f(x_2) = (1.7333)^4 - 12 = -2.974 < 0$

$\therefore$ Root lies between 1.7333 & 2. Taking, we get

$$\therefore x_3 = 1.7333 - \frac{2-1.7333}{4-(-2.974)}\times(-2.974) = 1.7333 + \frac{0.2667}{6.974}\times 2.974 = 1.8470$$

Now $f(x_3) = (1.847)^4 - 12 = -0.3623 < 0$.

$\therefore$ Root lies between 1.8470 & 2.

$$\therefore x_4 = 1.8470 - \frac{2-1.8470}{4-(-0.3623)}\times(-0.3623) = 1.8470 + \frac{0.153}{4.3623}\times 0.3623 = 1.8597$$

Now $f(x_4) = (1.8597)^4 - 12 = -0.0389 < 0$.

$\therefore$ Root lies between 1.8597& 2.

$$\therefore x_5 = 1.8597 - \frac{2-1.8597}{4-(-0.0389)}\times(-0.0389) = 1.8597 + \frac{0.1403}{4.0389}\times 0.0389 = 1.8611$$

$f(x_5) = -0.0028 < 0 \Rightarrow$ root lies between 1.8611 & 2

$$x_6 = 1.8611 - \frac{2-1.8611}{4-(-0.0028)}\times(-0.0028) = 1.8611 + \frac{0.1389}{4.0028}\times 0.0028 = 1.8612$$

$\therefore (12)^{1/4} = 1.861$ correct to 3 decimal places.

Exercises:

I Find the real root of the following equations correct to 3 decimal places by the method of false position.

A (1) $x^3 - 3x + 4 = 0$ $(Ans: -2.196)$ (2) $x^3 - 4x - 9 = 0$ $(Ans: 2.7065)$

(3) $x^3 + x - 1 = 0$ near $x = 1$ $(Ans: 0.682)$ (4) $x^6 - x^4 - x^3 - 1 = 0$ $(Ans: 1.404)$

B (1) $xe^x - 2 = 0$ $(Ans: 0.853)$ (2) $2x - \log x = 6$ $(Ans: 3.647)$

(3) $xe^x - \sin x = 0 \quad (Ans: -2.991)$

II Use the method of false position to find the fourth root of 32 correct to 3 decimal places. $(Ans: 2.378)$

III Find the root of the equation $xe^x = \cos x$ using regula falsi method correct to 4 decimal places. $(Ans: 0.5177)$

---

3 Newton – Raphson method

Let x_o be an approximation root of $f(x) = 0$. If $x_1 = x_o + h$ be the exact root, then $f(x_1) = 0$, *i.e.* $f(x_o + h) = 0$

Expanding $f(x_o + h)$ by using Taylor's series we get,

$$f(x_o) + hf^1(x_o) + \frac{h^2}{2!} f^{11}(x_o) + ... = 0$$

Since h is small, neglecting h^2 & higher powers of h, we get

$$f(x_o) + hf^1(x_o) = 0 \Rightarrow h = -\frac{f(x_o)}{f^1(x_o)}$$

$\therefore x_1 = x_o + h = x_o - \dfrac{f(x_o)}{f^1(x_o)}$, which is a closer approximation to the root.

Similarly starting with x_1, a still better approximation x_2 is given by $x_2 = x_1 - \dfrac{f(x_1)}{f^1(x_1)}$

In general $x_{n+1} = x_n - \dfrac{f(x_n)}{f^1(x_n)}$, which is known as the **Newton – Raphson formula** Or **Newton's iteration formula.**

$y = f(x)$

Geometrically

Let α be the root of the equation $f(x) = 0$.

Let x_o be a point near the root α. Then the

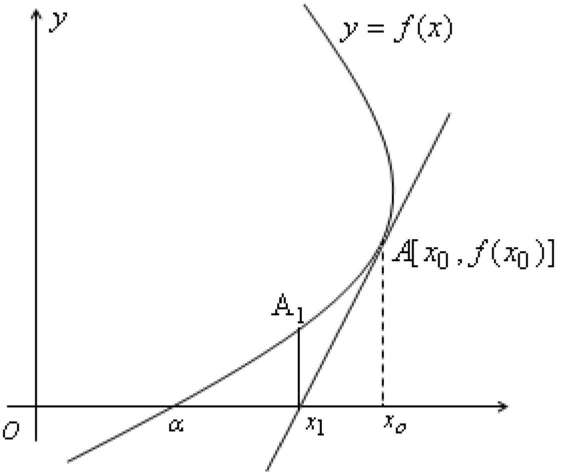
Equation of the tangent at $A[x_o, f(x_o)]$ is

$$f^1(x_0) = \frac{y - f(x_o)}{x - x_o}.$$

It cuts x – axis (*i.e.* $y = 0$) at

$$x_1 = x_o - \frac{f(x_o)}{f^1(x_o)} \quad [\text{take } x = x_1]$$

This is the first approximation to the root α.



Repeating this process, we approach to the root α. Hence this method consists in replacing the part of the curve between the point A & x – axis by means of the tangent to the curve at A.

Examples: **1)** Find by Newton's method, the real root of the equation $3x = \cos x + 1$

[**VTU 2003**]

Answer: Here $f(x) = 3x - \cos x - 1$

$\therefore f(0) = -2 < 0 \,\&\, f(1) = 1.4597 > 0 \Rightarrow$ a root lies between 0 & 1.

So take $x_o = 0.5$ [which is mid point of 0 & 1]

Now $f^1(x) = 3 + \sin x$

$\therefore$ By Newton's iteration formula we have,

$$x_{n+1} = x_n - \frac{f(x_n)}{f^1(x_n)} = x_n - \frac{3x_n - \cos x_n - 1}{3 + \sin x_n} = \frac{x_n \sin x_n + \cos x_n + 1}{3 + \sin x_n}$$

Putting $n = 0$, the first approximation x_1 is given by

$$x_1 = \frac{x_o \sin x_o + \cos x_o + 1}{3 + \sin x_o} = \frac{(0.5)\sin(0.5) + \cos(0.5) + 1}{3 + \sin(0.5)} = 0.6085$$

Putting $n = 1$, the second approximation x_2 is given by

$$x_2 = \frac{x_1 \sin x_1 + \cos x_1 + 1}{3 + \sin x_1} = \frac{(0.6085)\sin(0.6085) + \cos(0.6085) + 1}{3 + \sin(0.6085)} = 0.6071$$

Putting $n = 2$, the third approximation x_3 is given by

$$x_3 = \frac{x_2 \sin x_2 + \cos x_2 + 1}{3 + \sin x_2} = \frac{(0.6071)\sin(0.6071) + \cos(0.6071) + 1}{3 + \sin(0.6071)} = 0.6071$$

Here $x_2 = x_3 \Rightarrow$ the root is 0.6071, correct to 4 decimal places.

---

2) Find the real root of $x \log_{10} x = 1.2$ by using Newton – Raphson method, correct to five decimal places.

Answer: Here $f(x) = x \log_{10} x - 1.2$

$\therefore f(1) = -1.2 < 0, f(2) = 2\log_{10} 2 - 1.2 = -0.59794 < 0, f(3) = 3\log_{10} 3 - 1.2 = 0.23136 > 0$

So a root lies between 2 & 3, Let $x_o = 2.5$ be the initial approximation to the root.

Now $f^1(x) = \log_{10} x + x.\frac{1}{x}\log_{10} e = \log_{10} x + \log_{10} e$ $\qquad [\because \frac{d}{dx}(\log_a x) = \frac{1}{x}\log_a e]$

$= \log_{10} x + 0.43429 \qquad [\because \log_{10} e = 0.43429]$

Hence Newton – Raphson formula gives-

$$x_{n+1} = x_n - \frac{f(x_n)}{f^1(x_n)} = x_n - \frac{x_n \log_{10} x_n - 1.2}{\log_{10} x_n + 0.43429} = \frac{0.43429 x_n + 1.2}{\log_{10} x_n + 0.43429}$$

$$\therefore x_1 = \frac{0.43429x_o + 1.2}{\log_{10} x_o + 0.43429} = \frac{0.43429 \times 2.5 + 1.2}{\log_{10} 2.5 + 0.43429} = 2.746506$$

$$x_2 = \frac{0.43429x_1 + 1.2}{\log_{10} x_1 + 0.43429} = \frac{0.43429 \times 2.746506 + 1.2}{\log_{10} 2.746506 + 0.43429} = 2.740649$$

$$x_3 = \frac{0.43429x_2 + 1.2}{\log_{10} x_2 + 0.43429} = \frac{0.43429 \times 2.740649 + 1.2}{\log_{10} 2.740649 + 0.43429} = 2.740646$$

$\Rightarrow x = 2.74065$ is the required root correct to 5 decimal places.

---

3) Using N – R method, find a root of $\tan x = 1.5x$ which is near $x = 1$ correct to 3 decimal places.

Answer: Here $f(x) = \tan x - 1.5x \Rightarrow f^1(x) = \sec^2 x - 1.5$ & take $x_o = 1(given)$

Hence Newton – Raphson formula gives-

$$x_{n+1} = x_n - \frac{f(x_n)}{f^1(x_n)} = x_n - \frac{\tan x_n - 1.5x_n}{\sec^2 x_n - 1.5} = \frac{x_n \sec^2 x_n - \tan x_n}{\sec^2 x_n - 1.5}$$

$$= \frac{\sec^2 x_n (x_n - \sin x_n \cos x_n)}{\sec^2 x_n \left(1 - \frac{1.5}{\sec^2 x_n}\right)} = \frac{x_n - \frac{1}{2}\sin 2x_n}{1 - 1.5\cos^2 x_n}$$

$$\therefore \ x_1 = \frac{x_o - 0.5\sin 2x_0}{1 - 1.5\cos^2 x_0} = \frac{1 - 0.5\sin 2}{1 - 1.5\cos^2 1} = 0.9702$$

$$x_2 = \frac{0.9702 - 0.5\sin(2 \times 0.9702)}{1 - 1.5\cos^2 0.9702} = 0.9672$$

$$x_3 = \frac{0.9672 - 0.5\sin(2 \times 0.9672)}{1 - 1.5\cos^2 0.9672} = 0.9674$$

$\therefore$ root of the equation is 0.967 correct to 3 decimal places.

Exercises:

I. Find by Newton's method, a root of the following correct to 3 decimal places:

(1) $x^3 - 3x + 1 = 0$ $(Ans: 0.347)$ (2) $x^3 - 2x - 5 = 0$ $(Ans: 2.095)$

(3) $x^3 - 5x + 3 = 0$ $(Ans: 0.657)$ (4) $x^3 - 6x + 4 = 0$ $(Ans: 0.732)$

(5) $3x^3 - 9x^2 + 8 = 0$, lying between 1 & 2 $(Ans: 1.226)$

II. Find a root of the following equations correct to 3 significant figure using Newton's iterative method:

(1) $x^4 - 12x + 7 = 0$ $(Ans: 0.594)$ (2) $x^4 - x - 13 = 0$ $(Ans: 1.967)$

(3) $x^4 + x^3 - 7x^2 - x + 5 = 0$, lying between 2 & 3 $(Ans: 2.061)$

(4) $x^5 - 5x^2 + 3 = 0$ $(Ans: 0.822)$

III. Using N – R method find a root of the following equations correct to 3 decimal places.

1) $xe^x - 2 = 0$ 2) $3\sin x - 2x + 5 = 0,$ near 3 3) $x\sin x + \cos x = 0$ which is near $x = \pi$

[Answers: 1) 0.853 2) 2.883 3) 2.798]

IV. Find by Newton's method, the root of the equations:

1) $\cos x = xe^x$ (VTU 2003) 2) $\log x - \cos x = 0$ 3) $10^x + x - 4 = 0$

4) $x + \log_{10} x = 3.375$

Answers: 1) 0.518 2) 1.303 3) 0.539 4) 2.911

Useful Deductions (Recurrence formula)

(1) Iterative formula to find $\frac{1}{N}$ is $x_{n+1} = x_n(2 - Nx_n)$

(2) Iterative formula to find $\sqrt{N}$ is $x_{n+1} = \frac{1}{2}\left(x_n + \frac{N}{x_n}\right)$

(3) Iterative formula to find $\frac{1}{\sqrt{N}}$ is $x_{n+1} = \frac{1}{2}\left(x_n + \frac{1}{Nx_n}\right)$

(4) Iterative formula to find $\sqrt[k]{N}$ is $x_{n+1} = \frac{1}{k}\left((k-1)x_n + \frac{N}{x_n^{k-1}}\right)$

Proofs: **1)** To find $\frac{1}{N}$.

Let $x = \frac{1}{N}$. Then $\frac{1}{x} - N = 0$. Take $f(x) = \frac{1}{x} - N \Rightarrow f^1(x) = -\frac{1}{x^2}$.

Then Newton Raphson formula gives

$$x_{n+1} = x_n - \frac{f(x_n)}{f^1(x_n)} = x_n - \frac{\frac{1}{x_n} - N}{-\frac{1}{x_n^2}} = x_n + \left(\frac{1}{x_n} - N\right)x_n^2 = x_n + x_n - Nx_n^2 = 2x_n - Nx_n^2$$

Or $$\boxed{x_{n+1} = x_n(2 - Nx_n)}$$

2) To find $\sqrt{N}$

Let $x=\sqrt{N}, \therefore x^2-N=0$

Take $f(x)=x^2-N$, then $f^1(x)=2x$

$\therefore$ By Newton's formula, we get

$$x_{n+1}=x_n-\frac{f(x_n)}{f^1(x_n)}=x_n-\frac{x_n^2-N}{2x_n}=\frac{2x_n^2-x_n^2+N}{2x_n}=\frac{x_n^2-N}{2x_n}=\frac{1}{2}\left(x_n+\frac{N}{x_n}\right)$$

Or

$$\boxed{x_{n+1}=\frac{1}{2}\left(x_n+\frac{N}{x_n}\right)}$$

3) To find $\frac{1}{\sqrt{N}}$

Let $x=\frac{1}{\sqrt{N}}, \therefore x^2-\frac{1}{N}=0$

Take $f(x)=x^2-\frac{1}{N}$, then $f^1(x)=2x$

Then Newton Raphson formula gives

$$x_{n+1}=x_n-\frac{f(x_n)}{f^1(x_n)}=x_n-\frac{x_n^2-\frac{1}{N}}{2x_n}=\frac{2x_n^2-x_n^2+\frac{1}{N}}{2x_n}=\frac{x_n^2+\frac{1}{N}}{2x_n}=\frac{1}{2}\left(x_n+\frac{1}{Nx_n}\right)$$

Or

$$\boxed{x_{n+1}=\frac{1}{2}\left(x_n+\frac{1}{Nx_n}\right)}$$

4) To find $\sqrt[k]{N}$

Let $x=\sqrt[k]{N}, \therefore x^k-N=0$

Take $f(x)=x^k-N$, then $f^1(x)=kx^{k-1}$

Then by N – R formula we get,

$$x_{n+1}=x_n-\frac{f(x_n)}{f^1(x_n)}=x_n-\frac{x_n^k-N}{kx_n^{k-1}}=\frac{kx_n^k-x_n^k+N}{kx_n^{k-1}}=\frac{1}{k}\left((k-1)x_n+\frac{N}{x_n^{k-1}}\right)$$

Or

$$\boxed{x_{n+1}=\frac{1}{k}\left((k-1)x_n+\frac{N}{x_n^{k-1}}\right)}$$

Problems: Evaluate the following correct to 4 decimal places by Newton's iteration method

1) $\frac{1}{31}$

The iteration formula to find $\frac{1}{N}$ is $x_{n+1} = x_n(2 - Nx_n)$, here $N = 31$, then we get

$$x_{n+1} = x_n(2 - 31x_n)$$

An approximate value of $\frac{1}{31}$ is 0.03, $\therefore$ take $x_o = 0.03$, then we get,

$$x_1 = x_o(2 - 31x_o) = 0.03(2 - 31 \times 0.03) = 0.0321$$

$$x_2 = x_1(2 - 31x_1) = 0.0321(2 - 31 \times 0.0321) = 0.03226$$

$$x_3 = x_2(2 - 31x_2) = 0.032257(2 - 31 \times 0.032257) = 0.03226$$

$\therefore x_2 = x_3$ up to 4 decimal places, we have $\frac{1}{31} = 0.0323$

---

2) $\frac{1}{\sqrt{14}}$

The iteration formula to find $\frac{1}{\sqrt{N}}$ is $x_{n+1} = \frac{1}{2}\left(x_n + \frac{1}{Nx_n}\right)$, here $N = 14$.

$$\therefore x_{n+1} = \frac{1}{2}\left(x_n + \frac{1}{14x_n}\right)$$

An approximate value of $\frac{1}{\sqrt{14}} \approx \frac{1}{\sqrt{16}} = \frac{1}{4} = 0.25$. $\therefore$ take $x_o = 0.25$, then we get

$$x_1 = \frac{1}{2}\left(x_o + \frac{1}{14x_o}\right) = \frac{1}{2}\left(0.25 + \frac{1}{14 \times 0.25}\right) = 0.26786$$

$$x_2 = \frac{1}{2}\left(x_1 + \frac{1}{14x_1}\right) = \frac{1}{2}\left(0.26786 + \frac{1}{14 \times 0.26786}\right) = 0.26726$$

$$x_3 = \frac{1}{2}\left(x_2 + \frac{1}{14x_2}\right) = \frac{1}{2}\left(0.26726 + \frac{1}{14 \times 0.26726}\right) = 0.26726$$

$\therefore x_2 = x_3$ up to 4 decimal places $\Rightarrow \frac{1}{\sqrt{14}} = 0.2673$

---

3) $\sqrt{28}$

The iteration formula to find $\sqrt{N}$ is $x_{n+1} = \frac{1}{2}\left(x_n + \frac{N}{x_n}\right)$, here $N = 28$

$\therefore x_{n+1} = \frac{1}{2}\left(x_n + \frac{28}{x_n}\right)$, An approximate value of $\sqrt{28} \approx \sqrt{25} = 5$, $\therefore$ take $x_o = 5$, we get

$$\therefore x_1 = \frac{1}{2}\left(x_o + \frac{28}{x_o}\right) = \frac{1}{2}\left(5 + \frac{28}{5}\right) = 5.3, \quad x_2 = \frac{1}{2}\left(x_1 + \frac{28}{x_1}\right) = \frac{1}{2}\left(5.3 + \frac{28}{5.3}\right) = 5.29151$$

$$x_3 = \frac{1}{2}\left(x_2 + \frac{28}{x_2}\right) = \frac{1}{2}\left(5.29151 + \frac{28}{5.29151}\right) = 5.29150$$

$\therefore x_2 = x_3$ up to 4 decimal places, we have $\sqrt{28} = 5.2915$

---

4) $\sqrt[3]{24}$

Answer: The iteration formula to find $\sqrt[k]{N}$ is $x_{n+1} = \frac{1}{k}\left((k-1)x_n + \frac{N}{x_n^{k-1}}\right)$, here

$N = 24 \,\&\, k = 3$

$\therefore\ x_{n+1} = \frac{1}{3}\left(2x_n + \frac{24}{x_n^2}\right)$, An approximate value of $\sqrt[3]{24} \approx \sqrt[3]{27} = 3.$ $\therefore$ take $x_o = 3.$ Then we get,

$$x_1 = \frac{1}{3}\left(2x_o + \frac{24}{x_o^2}\right) = \frac{1}{3}\left(2\times 3 + \frac{24}{3^2}\right) = 2.88889,$$

$$x_2 == \frac{1}{3}\left(2\times 2.88889 + \frac{24}{(2.88889)^2}\right) = 2.88451$$

$$x_3 == \frac{1}{3}\left(2\times 2.88451 + \frac{24}{(2.88451)^2}\right) = 2.88450$$

Since $x_2 = x_3$ up to 4 decimal places, we have $\sqrt[3]{24} = 2.8845$

5) $(30)^{-\frac{1}{5}}$

Answer: The iteration formula to find $N^{\frac{1}{k}}$ is $x_{n+1} = \frac{1}{k}\left((k-1)x_n + \frac{N}{x_n^{k-1}}\right)$, here

$N = 30 \,\&\, k = -5$

$$\therefore\ x_{n+1} = -\frac{1}{5}\left(-6x_n + \frac{30}{x_n^{-6}}\right) = \frac{x_n}{5}\left(6 - 30x_n^5\right)$$

An approximate value of $(30)^{-\frac{1}{5}} \approx (32)^{-\frac{1}{5}} = \frac{1}{2} = 0.5.$ $\therefore$ take $x_o = 0.5.$ Then we get,

$$x_1 = \frac{x_o}{5}\left(6 - 30x_o^5\right) = \frac{0.5}{5}\left(6 - 30(0.5)^5\right) = 0.50625$$

$$x_2 = \frac{0.50625}{5}\left(6 - 30(0.50625)^5\right) = 0.50650,\ x_3 = \frac{0.5065}{5}\left(6 - 30(0.5065)^5\right) = 0.50650$$

Since $x_2 = x_3$ up to 4 decimal places, we have $(30)^{-\frac{1}{5}} = 0.5065$

Exercises:

1) Develop a recurrence formula for finding $\sqrt{N},$ using Newton's – Raphson method and hence compute $\sqrt{32}$ correct to 4 decimal places.

2) Find the cube root of 41, using N – R method.

3) Develop an algorithm using N – R method to find the fourth root of a positive number N and hence find $\sqrt[4]{32}$

4) Evaluate the following (correct to 3 decimal places) by using Newton's – Raphson method.

i) $\frac{1}{18}$ ii) $\frac{1}{\sqrt{15}}$ iii) $(28)^{-\frac{1}{4}}$

Answers: 1) $x_{n+1} = \frac{1}{2}\left(x_n + \frac{N}{x_n}\right)$, 5.6569 2) 3.4482 3) 2.3784

4) i) 0.056 ii) 0.258 iii) 0.4347

Finite differences

1) <u>Forward differences</u>: Let $y_0, y_1, y_2, \ldots\ldots.., y_n$ denote a set of values of $y = f(x)$ at the points $x_0, x_1, x_2, \ldots\ldots.., x_n$ respectively. Then the difference $y_n - y_{n-1}$ is called forward differences of y and is denoted by Δy_{n-1}.

$\therefore \Delta y_{n-1} = y_n - y_{n-1}$, $\Delta y_0 = y_1 - y_0$, $\Delta y_1 = y_2 - y_1$, $\Delta y_2 = y_3 - y_2$, ... , where Δ is called forward difference operator and Δy_0, Δy_1, $\Delta y_2, \ldots\ldots\ldots.$ are called 1^{st} forward differences. Similarly the differences of the 1^{st} forward differences are denoted by $\Delta^2 y_0, \Delta^2 y_1, \Delta^2 y_2, \ldots\ldots\ldots.$ & similarly we can define 3^{rd}, 4^{rd} forward differences etc.

Thus $\Delta^2 y_0 = \Delta y_1 - \Delta y_0 = y_2 - y_1 - (y_1 - y_0) = y_2 - 2y_1 - y_0$

$\Delta^2 y_1 = \Delta y_2 - \Delta y_1 = y_3 - y_2 - (y_2 - y_1) = y_3 - 2y_2 - y_1$

Similarly $\Delta^2 y_2 = y_4 - 2y_3 - y_2$, $\Delta^2 y_3 = y_5 - 2y_4 - y_3$ & soon

Now $\Delta^3 y_0 = \Delta^2 y_1 - \Delta^2 y_0 = y_3 - 2y_2 + y_1 - (y_2 - 2y_1 + y_0) = y_3 - 3y_2 + 3y_1 - y_0$

Similarly $\Delta^3 y_1 = y_4 - 3y_3 + 3y_2 - y_1$, $\Delta^3 y_2 = y_5 - 3y_4 + 3y_3 - y_2$ & soon.

In the similar manner we get, $\Delta^4 y_0 = y_4 - 4y_3 + 6y_2 - 4y_1 + y_0$,

$\Delta^5 y_0 = y_5 - 5y_4 + 10y_3 - 10y_2 + 5y_1 - y_0$,

$\Delta^6 y_0 = y_6 - 6y_5 + 15y_4 - 20y_3 + 15y_2 - 6y_1 + y_0$ & soon

Hence we can express any higher order differences in terms of the y. Since the coefficient on the right side are the binomial coefficients that can be obtained by **"Pascal's Triangle"** Which is as shown below

$n = 0$ 1

$n = 1$ 1 1

$n=2$ 1 2 1

$n=3$ 1 3 3 1

$n=4$ 1 4 6 4 1

$n=5$ 1 5 10 10 5 1

$n=6$ 1 6 15 20 15 6 1

$n=7$ 1 7 21 35 35 21 7 1 soon

The forward differences are usually arranged in tabular column, called a Forward difference table

Forward difference table:-

x	y	Δ	Δ^2	Δ^3	Δ^4
x_0	y_0				
		Δy_0			
x_1	y_1		$\Delta^2 y_0$		
		Δy_1		$\Delta^3 y_0$	
x_2	y_2		$\Delta^2 y_1$		$\Delta^4 y_0$
		Δy_2			
x_3	y_3		$\Delta^2 y_2$	$\Delta^3 y_1$	
		Δy_3			
x_4	y_4				

2. Backward differences: Let $y_0, y_1, y_2, \ldots\ldots\ldots, y_n$ denote a set of values of $y = f(x)$ at the points $x_0, x_1, x_2, \ldots\ldots\ldots, x_n$ respectively. Then the difference $y_n - y_{n-1}$ is called Backward differences of y and is denoted by ∇y_n

$\therefore \nabla y_1 = y_1 - y_0,\ \nabla y_2 = y_2 - y_1,\ \nabla y_3 = y_3 - y_2, \ldots\ldots, \nabla y_n = y_n - y_{n-1}$

where ∇ is called forward difference operator and $\nabla y_n, \nabla y_{n-1}, \nabla y_{n-2}, \ldots\ldots\ldots$ are called 1^{st} backward differences. Like forward differences we can also define higher order backward differences.

Thus $\nabla^2 y_2 = \nabla y_2 - \nabla y_1 = y_2 - y_1 - (y_1 - y_0) = y_2 - 2y_1 + y_0$

$\nabla^2 y_3 = \nabla y_3 - \nabla y_2 = y_3 - y_2 - (y_2 - y_1) = y_3 - 2y_2 + y_1$

Similarly $\nabla^2 y_4 = y_4 - 2y_3 + y_2$ soon.

Now $\nabla^3 y_3 = \nabla^2 y_3 - \nabla^2 y_2 = y_3 - 2y_2 + y_1 - (y_2 - 2y_1 + y_0) = y_3 - 3y_2 + 3y_1 - y_0$

$\nabla^3 y_4 = y_4 - 3y_3 + 3y_2 - y_1$, & so on.

Backward difference table:-

x	y	∇	∇^2	∇^3	∇^4
x_0	y_0				
		∇y_1			
x_1	y_1		$\nabla^2 y_2$		
		∇y_2		$\nabla^3 y_3$	
x_2	y_2		$\nabla^2 y_3$		$\nabla^4 y_4$
		∇y_3		$\nabla^3 y_4$	
x_3	y_3		$\nabla^2 y_4$		
		∇y_4			
x_4	y_4				

Symbolic relations: In addition to the operation $\Delta \,\&\, \nabla$ we now define one more operator called Shift operator E as follows $Ey_r = y_{r+1}$, then

$E^2 y_r = Ey_{r+1} + y_{r+2}$ & soon. In general $E^n y_r = y_{r+n}$

Relation between E & Δ:

We know that $\Delta y_0 = y_1 - y_0 = Ey_0 - y_0 = (E-1)y_0 \Rightarrow \Delta = E-1$ or $E = \Delta + 1$ -------- (1)

Equation (1) ⇒ that the effect of the operator E on y_0 is the same as that of the operator $\Delta + 1$ on y_0 .

Now $\Delta^2 y_0 = (E-1)^2 y_0 = (E^2 - 2E + 1)y_0 = E^2 y_0 - 2Ey_0 + y_0 = y_2 - 2y_1 + y_0$

Relation between ∇ & E:

We know that $\nabla y_1 = y_1 - y_0 = y_1 - E^{-1}y_1 = (1-E^{-1})y_1 \Rightarrow \nabla = 1 - E^{-1}$ ------ (2)

Relation between Δ, ∇ & E:

$\nabla E = (1-E^{-1})E = E - 1 = \Delta \Rightarrow \Delta = \nabla E$ ------ (3)

Relation between Δ & ∇:

$$(\Delta\nabla)y_r = \Delta(\nabla y_r) = \Delta(y_r - y_{r-1}) = \Delta y_r - \Delta y_{r-1} = \Delta y_r - (y_r - y_{r-1}) = \Delta y_r - \nabla y_r = (\Delta - \nabla)y_r$$

$\therefore \Delta\nabla = \Delta - \nabla$ ----------- (4)

Examples:

1) The following Table gives a set of values of x and the corresponding values of $y = f(x)$:

x	10	15	20	25	30	35
$y = f(x)$	19.97	21.51	22.47	23.52	24.65	25.89

Form the difference table and write down the values of $\Delta f(10)$, $\Delta^2 f(10)$, $\Delta^3 f(20)$, $\Delta^4 f(15)$, $\nabla f(20)$, $\nabla^2 f(30)$, $\nabla^3 f(25)$, and $\nabla^4 f(35)$

The difference table is

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	$\Delta^5 y$

10	19.97					
		1.54				
15	21.51		-0.58			
		0.96		0.67		
20	22.47		0.09		-0.68	
		1.05		-0.01		0.72
25	23.52		0.08		0.04	
		1.13		0.03		
30	24.65		0.11			
		1.24				
35	25.89					

$\therefore \Delta f(10) = \Delta y_0 = 1.54$, $\Delta^2 f(10) = \Delta^2 y_0 = -0.58,$ $\Delta^3 f(20) = \Delta^3 y2 = 0.03$, $\Delta^4 f(15) = \Delta^4 y_1 = 0.04$

$\nabla f(20) = \nabla y_2 = 0.96,$ $\nabla^2 f(30) = \nabla^2 y_4 = 0.08,$ $\nabla^3 f(25) = \nabla^3 y2 = 0.67,$ $\nabla^4 f(35) = \nabla^4 y_5 = 0.04$

Interpolation

Introduction: Let $y = f(x)$ be a continuous function defined in some interval $[x_0, x_n]$, then we can find the functional values $f(x)$ corresponding to given values of x say $x_0, x_1, \ldots, x_n$ as $y_n = f(x_n)$.

Now for a given set of tabular values $(x_0, y_0), (x_1, y_1), \ldots\ldots, (x_n, y_n)$ satisfying the relation $y = f(x)$ where the explicit nature of $f(x)$ is not known, we can find a simpler function, say $\phi(x)$ such that $f(x)$ & $\phi(x)$ agree at the set of tabulated points and this process is called '**interpolation**'. The process of estimating the value of y for some value of x out side the interval (x_0, x_n) is called extrapolation. If $\phi(x)$ is a polynomial, then process is called polynomial interpolation & $\phi(x)$ is called the interpolating polynomial.

<u>Interpolation with equal interval</u>:

<u>Newton's interpolation formulae (Gregory – Newton)</u>:

Let $(x_0, y_0), (x_1, y_1), \ldots\ldots, (x_n, y_n)$ be a given set of $(n+1)$ values of x & y where values of x be equidistance. *i.e.* $x_i = x_0 + ih,\ i = 0, 1, 2, 3, \ldots\ldots\ldots, n$

$$y_n(x) = y_0 + \frac{p\Delta y_0}{1!} + \frac{p(p-1)\Delta^2 y_0}{2!} + \frac{p(p-1)(p-2)\Delta^3 y_0}{3!} + \ldots\ldots + \frac{p(p-1)(p-2)\ldots\ldots(p-n+1)\Delta^n y_0}{n!}$$

Where $p=\dfrac{x-x_0}{h}$ This is called as **Newton's or Newton-Gregory forward difference interpolation formula** & is used to determine the interpolation near the beginning of a set of tabular values &

$$y_n(x)=y_n+\frac{p\nabla y_n}{1!}+\frac{p(p+1)\nabla^2 y_n}{2!}+\frac{p(p+1)(p+2)\nabla^3 y_n}{3!}+.........+\frac{p(p+1)(p+2)..........(p+n-1)\nabla^n y}{n!}$$

Where $p=\dfrac{x-x_n}{h}$. This is called **Newton's or Newton-Gregory backward difference interpolation formula** & is used to interpolating near the end of the tabulated values.

Problems:

1) Find the interpolation Polynomial which takes the following values. $y(0)=1,\ y(1)=0,$ $y(2)=1,\ \&\ y(3)=10$ and hence find $y(0.5),\ y(4)\ \&\ y(5)$

The difference table is

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$
0	1			
		-1		
1	0		2	
		1		6
2	1		8	
		9		
3	10			

Using forward formula we get;

Here $h=1,\ x_0=0\ \therefore p=\dfrac{x-x_0}{h}=\dfrac{x-0}{1}=x$

Newton's forward formula is

$$y_n(x)=y_0+\frac{p\Delta y_0}{1!}+\frac{p(p-1)\Delta^2 y_0}{2!}+\frac{p(p-1)(p-2)\Delta^3 y_0}{3!}$$

$$=1+\frac{x\times(-1)}{1}+\frac{x(x-1)\times 2}{2}+\frac{x(x-1)(x-2)\times 6}{6}$$

$$=1-x+x^2-x+x^3-3x^2+2x=x^3-2x^2+1$$

Hence $y(0.5)=(0.5)^3-2(0.5)^2+1=0.125-2\times 0.25+1=0.125-0.5+1=0.625$

$y(4)=4^3-2\times 4^2+1=64-32+1=33$ & $y(5)=5^3-2\times 5^2+1=125-50+1=76$

Using backward formula we get;

Here $h=1,\ x_n=3.\ \therefore p=\dfrac{x-x_n}{h}=\dfrac{x-3}{1}=x-3$

Newton's backward formula is

$$y_n(x)=y_n+\frac{p\nabla y_n}{1!}+\frac{p(p+1)\nabla^2 y_n}{2!}+\frac{p(p+1)(p+2)\nabla^3 y_n}{3!}$$

$$= 10+\frac{(x-3)\times 9}{1}+\frac{(x-3)(x-2)\times 8}{2}+\frac{(x-3)(x-2)(x-1)\times 6}{6}$$

$$=10+9x-27+4x^2-20x+24+x^3-6x^2+11x-6=x^3-2x^2+1$$

Note: The above example shows that if a tabulated function is a polynomial, then forward & backward formula gives same results.

2) Given that $\sin 45^o = 0.7071,\ \sin 50^o = 0.7660,\ \sin 55^o = 0.8192,\ \sin 60^o = 0.8660$, find $\sin 48^o$.

Answer: The Given data can put in tabular form as –

x (in degrees)	45	50	55	60
$f(x)=\sin x$	0.7071	0.7660	0.8192	0.8660

The difference Table is

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$
45	0.7071			
		0.0589		
50	0.7660		-0.0057	
		0.0532		-0.0007
55	0.8192		-0.0064	
		0.0468		
60	0.8660			

Here $h=5,\ x_0=45,\ x=48 .\therefore\ p=\frac{x-x_o}{h}=\frac{48-45}{5}=\frac{3}{5}=0.6$

Consider forward difference formula,

$$y_n(x)=y_0+\frac{p\Delta y_0}{1!}+\frac{p(p-1)\Delta^2 y_0}{2!}+\frac{p(p-1)(p-2)\Delta^3 y_0}{3!}$$

$$=0.7071+\frac{0.6\times 0.0589}{1}+\frac{0.6(-0.4)\times(-0.0057)}{2}+\frac{0.6(-0.4)(-1.4)\times(-0.0007)}{6}=0.7430848$$

$\therefore \sin 48^o = 0.7431$, correct to 4 decimal places.

3) The following table gives the distances in miles of the visible horizon for the given heights in feet above the earth's surface. Find y for $x=218$

$x:$	200	250	300	350	400
$f(x):$	15.04	16.81	18.42	19.9	21.27

Answer: The difference table is

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
200	15.04				
		1.77			
250	16.81		-0.16		
		1.61		0.03	
300	18.42		-0.13		-0.01
		1.48		0.02	
350	19.9		-0.11		
		1.37			
400	21.27				

Here $h=5,\ x_0=200,\ x=218 . \therefore\ p=\frac{x-x_o}{h}=\frac{218-200}{50}=\frac{18}{50}=0.36$

Consider forward difference formula,

$$y_n(x)=y_0+\frac{p\Delta y_0}{1!}+\frac{p(p-1)\Delta^2 y_0}{2!}+\frac{p(p-1)(p-2)\Delta^3 y_0}{3!}+\frac{p(p-1)(p-2)(p-3)\Delta^4 y_0}{4!}$$

$$=15.04+\frac{0.36\times 1.77}{1}+\frac{0.36(-0.64)\times(-0.16)}{2}+\frac{0.6(-0.64)(-1.64)\times 0.03}{6}+\frac{0.6(-0.64)(-1.64)(-2.64)\times(-0.01}{24}$$

$=15.697936921$

$\therefore\ y=15.6979$, correct to 4 decimal places

4) From the following table, estimate the number of students who obtained marks between 40 and 45:

Marks	30 - 40	40 - 50	50 - 60	60 - 70	70 - 80
No. of students	31	42	51	35	31

Answer: The cumulative frequency table is

Marks less than : x	40	50	60	70	80
No. of students: y	31	73	124	159	190

Then the difference table is

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
40	31				
		42			
50	73		9		
		51		-25	
60	124		-16		37
		35		12	
70	259		-4		
		31			
80	190				

Here $h=10,\ x_o=40,\ x=45$ (*i.e.* we have to find y when x = 45)

$$\therefore p = \frac{x-x_o}{h} = \frac{45-40}{10} = 0.5$$

Consider forward difference formula,

$$y_n(x) = y_0 + \frac{p\Delta y_0}{1!} + \frac{p(p-1)\Delta^2 y_0}{2!} + \frac{p(p-1)(p-2)\Delta^3 y_0}{3!} + \frac{p(p-1)(p-2)(p-3)\Delta^4 y_0}{4!}$$

$$= 31 + \frac{0.5}{1}\times 42 + \frac{0.5(-0.5)}{2}\times 9 + \frac{0.5(-0.5)(-1.5)}{6}\times(-25) + \frac{0.5(-0.5)(-1.5)(-2.5)}{24}\times 37$$

$$= 47.8671875$$

$\therefore$ The number of student with marks less than 45 is $47.86 \approx 48$

But the number of students with marks less than 40 is 31.

Hence the number of students getting marks between 4 & 45 = 48 – 31=17.

5) In the table below, the values of y are consecutive terms of a series of which 23.6 is the 6^{th} term. Find the

first and tenth terms of the series.

x :	3	4	5	6	7	8	9
y:	4.8	8.4	14.5	23.6	36.2	52.8	73.9

Answer: The difference table is

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
3	4.8				
		3.6			
4	8.4		2.5		
		6.1		0.5	
5	14.5		3.0		0
		9.1		0.5	
6	23.6		3.5		0
		12.6		0.5	
7	36.2		4.0		0
		16.6		0.5	
8	52.8		4.5		
		21.1			
9	73.9				

To find first term, consider Newton's forward interpolation formula

Here $h = 1,\ x_o = 3,\ x = 1 \Rightarrow p = \frac{x-x_o}{h} = \frac{1-3}{1} = -2$

$$y_n(x) = y_0 + \frac{p\Delta y_0}{1!} + \frac{p(p-1)\Delta^2 y_0}{2!} + \frac{p(p-1)(p-2)\Delta^3 y_0}{3!} + \frac{p(p-1)(p-2)(p-3)\Delta^4 y_0}{4!}$$

$$= 4.8 + \frac{-2}{1}\times 3.6 + \frac{(-2)(-3)}{2}\times 2.5 + \frac{(-2)(-3)(-4)}{6}\times 0.5 + \frac{(-2)(-3)(-4)(-5)}{24}\times 0 = 3.1$$

$\therefore$ The first term is 3.1

To find tenth term, consider Newton's backward interpolation formula

Here $h=1,\ x_n=9,\ x=10 \Rightarrow p=\frac{x-x_n}{h}=\frac{10-9}{1}=1$

$$y_n(x)=y_n+\frac{p\nabla y_n}{1!}+\frac{p(p+1)\nabla^2 y_n}{2!}+\frac{p(p+1)(p+2)\nabla^3 y_n}{3!}+\frac{p(p+1)(p+2)(p+3)\nabla^4 y_n}{4!}$$

$$=73.9+\frac{1}{1}\times 21.1+\frac{1\times 2}{2}\times 4.5+\frac{1\times 2\times 3}{6}\times 0.5+\frac{1\pm 2\times 3\times 4}{24}\times 0=100$$

$\therefore$ The tenth term is 100.

6) Estimate the value of $f(22)$ & $f(42)$ from the following available data:

x	20	25	30	35	40	45
f(x)	354	332	291	260	231	204

Answer: The difference table is

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	$\Delta^5 y$
20	354					
		-22				
25	332		-19			
		-41		29		
30	291		10		-37	
		-31		-8		45
35	260		2		8	
		-29		0		
40	231		2			
		-27				
45	204					

To find $f(22)$ use Forward formula

Here $h=5,\ x_o=20\ \&\ x=22 \Rightarrow p=\frac{x-x_o}{h}=\frac{22-20}{5}=0.4$

Newton's forward interpolation formula is

$$y_n(x)=y_o+\frac{p\Delta y_o}{1!}+\frac{p(p-1)\Delta^2 y_o}{2!}+\frac{p(p-1)(p-2)\Delta^3 y_o}{3!}+\frac{p(p-1)(p-2)(p-3)\Delta^4 y_o}{4!}$$
$$+\frac{p(p-1)(p-2)(p-3)(p-4)\Delta^5 y_o}{5!}$$

$$=354+\frac{0.4}{1}\times(-22)+\frac{0.4\times(-0.6)}{2}\times(-19)+\frac{0.4\times(-0.6)(-1.6)}{6}\times 29+\frac{0.4\times(-0.6)(-1.6)(-2.6)}{24}\times(-37)$$
$$+\frac{0.4\times(-0.6)(-1.6)(-2.6)(-3.6)}{120}\times 45=351.70464$$

$\therefore\ f(22)=352$ approximately.

To find $f(42)$ use backward formula

Here $h=5,\ x_n=45\ \&\ x=42 \Rightarrow p=\dfrac{x-x_n}{h}=\dfrac{42-45}{5}=-0.6$

Newton's backward interpolation formula is

$$y_n(x)=y_n+\frac{p\nabla y_n}{1!}+\frac{p(p+1)\nabla^2 y_n}{2!}+\frac{p(p+1)(p+2)\nabla^3 y_n}{3!}+\frac{p(p+1)(p+2)(p+3)\nabla^4 y_n}{4!}$$
$$+\frac{p(p+1)(p+2)(p+3)(p+4)\nabla^5 y_n}{5!}$$

$$=204+\frac{-0.6}{1}\times(-27)+\frac{-0.6\times(0.4)}{2}\times 2+\frac{-0.6\times(0.4)(1.4)}{6}\times 0+\frac{-0.6\times(0.4)(1.4)(2.4)}{24}\times 8$$
$$+\frac{-0.6\times(0.4)(1.4)(2.4)(3.4)}{120}\times 45=218.66304$$

$\therefore\ f(42)=219$ approximately.

Exercises:

1) Find the cubic polynomial which takes the following values:

x:	0	1	2	3
f(x):	1	2	1	10

Hence find $f(4)$.

2) If $f(1.15)=1.0723$, $f(1.20)=1.0954$, $f(1.25)=1.1180$ and $f(1.30)=1.1401$, find $f(1.28)$

3) The area A of a circle of diameter d is given for the following values:

d:	80	85	90	95	100
A:	5026	5674	6362	7088	7854

Calculate the area of a circle of diameter 105

4) Find the number of men getting wages between Rs. 10 and 15 from the following data:

Wages in Rs. :	0 - 10	10 - 20	20 - 30	30 - 40
Frequency :	9	30	35	42

5) From the following data estimate the number of persons having incomes between 2000 and 2500

Income :	Below 500	500 - 1000	1000 - 2000	2000 - 3000	3000 - 4000
No. of persons :	6000	4250	3600	1500	650

6) In the following table, the values of y are consecutive terms of a series of which 12.5 is the 5^{th} term. Find the first and tenth terms of the series.

x :	3	4	5	6	7	8	9

y :	2.7	6.4	12.5	21.6	34.3	51.2	72.9

7) Find the cubic polynomial which takes the following values:

x :	0	1	2	3	4
f(x) :	-5	-10	-9	4	35

8) Given $\log_{10} 654 = 2.8156$, $\log_{10} 656 = 2.8169$, $\log_{10} 658 = 2.8182$, $\log_{10} 660 = 2.8195$, $\log_{10} 662 = 2.821$, find $\log_{10} 655$

9) From the data given in the following table, find the number of workers whose hourly wage is between Rs. 21 & Rs. 25.

Hourly wage (in Rs.):	0 - 10	11 - 20	21 - 30	31 - 40	41 - 50
No. of workers:	20	45	115	210	115

10) Estimate the value of $\tan(0.12)$ & $\tan(0.27)$ from the table given below:

x :	0.10	0.15	0.20	0.25	0.30
$y = \tan x$	0.1003	0.1511	0.2027	0.2553	0.3093

Answer: 1) $2x^3 - 7x^2 + 6x + 1$, $f(4) = 41$ 2) 1.1312 3) 8666 4) 24 5) 1169

6) 0.1 & 100 7) $x^3 - 6x - 5$ 8) 2.8162 9) 42 10) 0.1205, 0.2767

Interpolation with unequally spaced points:

In the Newton's interpolation we required the values of the independent variable to be equally spaced.

$\therefore$ Now we have to derive interpolation formula with unequally spaced values of the argument.

Here we have Lagrange's formula.

Lagrange's interpolation formula:

Let $y(x)$ be a continuous & differentiable $(n+1)$ times in the interval (a,b). Let (x_o, y_o), (x_1, y_1), ------, (x_n, y_n), be the set of $(n+1)$ points where the values of x need not necessarily be equally spaced. Then

$$f(x) = \frac{(x-x_1)(x-x_2).........(x-x_n)}{(x_o-x_1)(x_o-x_2).......(x_o-x_n)} \times y_o + \frac{(x-x_o)(x-x_2).........(x-x_n)}{(x_1-x_o)(x_1-x_2).......(x_1-x_n)} \times y_1 + ...$$

$$+\frac{(x-x_0)(x-x_1)\ldots\ldots\ldots(x-x_{n-1})}{(x_n-x_o)(x_n-x_1)\ldots\ldots\ldots(x_n-x_{n-1})}\times y_n \text{ -------(3)}$$

This is called **Lagrange's interpolation formula.**

Problems: 1) By using the Lagrange's interpolation formula, find $f(11)$ from the following data

x	2	5	8	14
$y=f(x)$	94.8	87.9	81.3	68.7

Answer: Here $x_0=2,\ x_1=5,\ x_2=8,\ x_3=14,\ y_0=94.8,\ y_1=87.9,\ y_2=81.3,\ y_3=68.7,$ and $x=11$

Lagrange's interpolation formula is

$$f(x)=\frac{(x-x_1)(x-x_2)(x-x_3)}{(x_o-x_1)(x_o-x_2)(x_o-x_3)}\times y_o+\frac{(x-x_o)(x-x_2)(x-x_3)}{(x_1-x_o)(x_1-x_2)(x_1-x_3)}\times y_1$$

$$+\frac{(x-x_o)(x-x_1)(x-x_3)}{(x_2-x_o)(x_2-x_1)(x_2-x_3)}\times y_2+\frac{(x-x_o)(x-x_1)(x-x_2)}{(x_3-x_o)(x_3-x_1)(x_3-x_2)}\times y_3$$

$\therefore$

$$f(11)=\frac{(11-5)(11-8)(11-14)}{(2-5)(2-8)(2-14)}\times 94.8+\frac{(11-2)(11-8)(11-14)}{(5-2)(5-8)(5-14)}\times 87.9+\frac{(11-2)(11-5)(11-14)}{(8-2)(8-5)(8-14)}\times 81.3$$

$$+\frac{(11-2)(11-5)(11-8)}{(14-2)(14-5)(14-8)}\times 68.7$$

$$=\frac{6\times 3\times(-3)}{(-3)(-6)(-12)}\times 94.8+\frac{9\times 3\times(-3)}{3\times(-3)(-9)}\times 87.9+\frac{9\times 6\times(-3)}{6\times 3\times(-6)}\times 81.3+\frac{9\times 6\times 3}{12\times 9\times 6}\times 68.7=74.925$$

2) Certain corresponding values of x & $\log_{10}x$ are: $(300,\ 2.4771)$, $(304,\ 2.4829)$ $(305,\ 2.4843)$ and $(307,\ 2.4871)$. Find $\log_{10}301$ by using the Lagrange's interpolation formula

Answer: Here $x_o=300,\ x_1=304,\ x_2=305,\ x_3=307,\ y_o=2.4771,\ y_1=2.4829,$ $y_2=2.4843,\ y_3=2.4871,$

and $x=301$

Lagrange's interpolation formula is

$$f(x)=\frac{(x-x_1)(x-x_2)(x-x_3)}{(x_o-x_1)(x_o-x_2)(x_o-x_3)}\times y_o+\frac{(x-x_o)(x-x_2)(x-x_3)}{(x_1-x_o)(x_1-x_2)(x_1-x_3)}\times y_1$$

$$+\frac{(x-x_o)(x-x_1)(x-x_3)}{(x_2-x_o)(x_2-x_1)(x_2-x_3)}\times y_2+\frac{(x-x_o)(x-x_1)(x-x_2)}{(x_3-x_o)(x_3-x_1)(x_3-x_2)}\times y_3$$

$$\therefore \log_{10} 301=\frac{(301-304)(301-305)(301-307)}{(300-304)(300-305)(300-307)}\times 2.4771+\frac{(301-300)(301-305)(301-307)}{(304-300)(304-305)(304-307)}\times 2.4829$$

$$+\frac{(301-300)(301-304)(301-307)}{(305-300)(305-304)(305-307)}\times 2.4843+\frac{(301-300)(301-304)(301-305)}{(307-300)(307-304)(307-305)}\times 2.4871$$

$$=\frac{(-3)(-4)(-6)}{(-4)(-5)(-7)}\times 2.4771+\frac{1\times(-4)(-6)}{4\times(-1)(-3)}\times 2.4829+\frac{1\times(-3)(-6)}{5\times 1\times(-2)}\times 2.4843+\frac{1\times(-3)(-4)}{7\times 3\times 2}\times 2.4871$$

$$=2.4786$$

3) Given the values

x	5	7	11	13	17
$f(x)$	150	392	1452	2366	5202

Estimate $f(9)$ using Lagrange's interpolation formula

Answer: Here $x_o=5,\ x_1=7,\ x_2=11,\ x_3=13,\ x_4=17\ y_o=150,\ y_1=392,\ y_2=1452,$

$y_3=2366,\ y_4=5202$

and $x=9$

Lagrange's interpolation formula is

$$f(x)=\frac{(x-x_1)(x-x_2)(x-x_3)(x-x_4)}{(x_o-x_1)(x_o-x_2)(x_o-x_3)(x_0-x_4)}\times y_o+\frac{(x-x_o)(x-x_2)(x-x_3)(x-x_4)}{(x_1-x_o)(x_1-x_2)(x_1-x_3)(x_1-x_4)}\times y_1$$

$$+\frac{(x-x_o)(x-x_1)(x-x_3)(x-x_4)}{(x_2-x_o)(x_2-x_1)(x_2-x_3)(x_2-x_4)}\times y_2+\frac{(x-x_o)(x-x_1)(x-x_2)(x-x_4)}{(x_3-x_o)(x_3-x_1)(x_3-x_2)(x_3-x_4)}\times y_3$$

$$+\frac{(x-x_o)(x-x_1)(x-x_2)(x-x_3)}{(x_4-x_o)(x_4-x_1)(x_4-x_2)(x_4-x_3)}\times y_4$$

$$\therefore\ f(9)=\frac{(9-7)(9-11)(9-13)(9-17)}{(5-7)(5-11)(5-13)(5-17)}\times 150+\frac{(9-5)(9-11)(9-13)(9-17)}{(7-5)(7-11)(7-13)(7-17)}\times 392$$

$$+\frac{(9-5)(9-7)(9-13)(9-17)}{(11-5)(11-7)(11-13)(11-17)}\times 1452+\frac{(9-5)(9-7)(9-11)(9-17)}{(13-5)(13-7)(13-11)(13-17)}\times 2366$$

$$+\frac{(9-5)(9-7)(9-11)(9-13)}{(17-5)(17-7)(17-11)(17-13)}\times 5202$$

$$=\frac{2\times(-2)(-4)(-8)}{(-2)(-6)(-8)(-12)}\times 150+\frac{4\times(-2)(-4)(-8)}{2\times(-4)(-6)(-10)}\times 392+\frac{4\times 2(-4)(-8)}{6\times 4\times(-2)(-6)}\times 1452$$

$$+\frac{4\times 2(-2)(-8)}{8\times 6\times 2(-4)}\times 2366\ +\frac{4\times 2(-2)(-4)}{12\times 10\times 6\times 4}\times 5202$$

$$=809.99999978\approx 810$$

4) Using Lagrange's interpolation formula, find y when $x=10$ from the following table:

x	5	6	9	11
y	12	13	14	16

Answer: Here $x_o=5,\ x_1=6,\ x_2=9,\ x_3=11,\ y_o=12,\ y_1=13,\ y_2=14,\ y_3=16$, and

$x=10$

Lagrange's interpolation formula is

$$f(x)=\frac{(x-x_1)(x-x_2)(x-x_3)}{(x_o-x_1)(x_o-x_2)(x_o-x_3)}\times y_o+\frac{(x-x_o)(x-x_2)(x-x_3)}{(x_1-x_o)(x_1-x_2)(x_1-x_3)}\times y_1$$

$$+\frac{(x-x_o)(x-x_1)(x-x_3)}{(x_2-x_o)(x_2-x_1)(x_2-x_3)}\times y_2+\frac{(x-x_o)(x-x_1)(x-x_2)}{(x_3-x_o)(x_3-x_1)(x_3-x_2)}\times y_3$$

$$\therefore\ f(10)=\frac{(10-6)(10-9)(10-11)}{(5-6)(5-9)(5-11)}\times 12+\frac{(10-5)(10-9)(10-11)}{(6-5)(6-9)(6-11)}\times 13$$

$$+\frac{(10-5)(10-6)(10-11)}{(9-5)(9-6)(9-11)}\times 14$$

$$+\frac{(10-5)(10-6)(10-9)}{(11-5)(11-6)(11-9)}\times 16$$

$$=\frac{4\times 1(-1)}{(-1)(-4)(-6)}\times 12+\frac{5\times 1(-1)}{1\times(-3)(-5)}\times 13+\frac{5\times 4(-1)}{4\times 3\times(-2)}\times 14+\frac{5\times 4\times 1}{6\times 5\times 2}\times 16$$

$$=14.66666\approx 14.67$$

Exercises:

1) Using Lagrange's interpolation formula, find $f(5)$ from the following data

x	1	3	4	6	9
$f(x)$	-3	9	30	132	156

2) Use Lagrange's interpolation formula to find $f(4)$. Given

x	0	2	3	6
$f(x)$	-4	2	14	158

3) If $y(1)=-3,\ y(3)=9,\ y(4)=30,\ y(6)=132,$ find the Lagrange's interpolation Polynomial that takes the

same values as y at the given points.

4) Given $u_o=707,\ u_1=819,\ u_3=866\ \&\ u_6=966,$ find u_4 using Lagrange's interpolation formula.

5) Using Lagrange's interpolation formula Find $f(6)$. Given

x	3	7	9	10
$f(x)$	168	120	72	63

Answers: 1) 73.1 2) 40 3) x^3-3x^2+5x-6 4) 906 5) 147

Newton's General Interpolation Formula (Divided – difference)

The Lagrange's interpolation formula, derived earlier has the disadvantage that if another interpolation point were added, then we have to change the Lagrange's interpolation formula completely. Therefore we seek an interpolation polynomial which has the property that a polynomial of higher degree may be derived from it by simply adding new terms. Newton's general interpolation formula is one such formula for which we need **divided differences**.

Let $(x_0,y_0),(x_1,y_1),\ldots\ldots\ldots,(x_n,y_n)$ be the given $(n+1)$ points. Then the divided differences of orders $1,\ 2,\ldots\ldots\ldots\ldots,\ n$ are defined as

$$[x_o,x_1]=\frac{y_1-y_o}{x_1-x_o},\ [x_o,x_1,x_2]=\frac{[x_1,x_2]-[x_o,x_1]}{x_2-x_o},$$

$$[x_o,x_1,x_2,x_3]=\frac{[x_1,x_2,x_3]-[x_o,x_1,{}_2\,]}{x_3-x_o},\ldots\ldots$$

$$\ldots\ldots\ldots\ldots\ldots\ldots,\ [x_o,x_1,x_2,\ldots\ldots,\ x_n]=\frac{[x_1,x_2,\ldots\ldots x_n]-[x_o,x_1,\ldots\ldots x_{n-1}]}{x_n-x_o}.$$

Note: $[x_o,x_1]=\dfrac{y_1-y_o}{x_1-x_o}=\dfrac{y_o-y_1}{x_o-x_1}=[x_1,x_o]$

Newton's divided-difference interpolation formula

From the definition of divided differences we have

$[x,x_o]=\dfrac{y-y_o}{x-x_o}\Rightarrow y=y_o+(x-x_o)[x,x_o]------(1)$ Continuing in this way we get

$$y=y_o+(x-x_o)[x_o,x_1]+(x-x_o)(x-x_1)[x_o,x_1,x_2]+(x-x_o)(x-x_1)(x-x_2)[x_o,x_1,x_2,x_3]$$

$$+ \ldots\ldots\ \ldots\ldots + (x - x_o)(x - x_1)(x - x_2)\ldots\ldots(x - x_n)[x, x_o, x_1, \ldots\ldots, x_n]$$

This is called **Newton's general interpolation formula** with divided differences, the last term is called the remainder term after (n+1) terms.

Problems: 1) Determine $f(x)$ as a polynomial in x for the following data, using Newton's divided difference Formula

x	- 4	-1	0	2	5
$f(x)$	1245	33	5	9	1335

Answer: The divided differences table is

x	y	1^{st} divided differences	2^{nd} divided differences	3^{rd} divided differences	4^{th} divided differences
- 4	**1245**				
		-404			
-1	33		**94**		
		-28		**-14**	
0	5		10		**3**
		2		13	
2	9		88		
		442			
5	1335				

Newton's divided differences interpolation formula is

$$y = y_o + (x - x_o)[x_o, x_1] + (x - x_o)(x - x_1)[x_o, x_1, x_2] + (x - x_o)(x - x_1)(x - x_2)[x_o, x_1, x_2, x_3]$$

$$+ (x - x_o)(x - x_1)(x - x_2)(x - x_3)[x_o, x_1, x_2, x_3, x_4].$$

Here $[x_o, x_1] = -404$, $[x_o, x_1, x_2] = 94$, $[x_o, x_1, x_2, x_3] = -14$, $[x_o, x_1, x_2, x_3, x_4] = 3$,

$$\therefore\ f(x) = 1245 + (x+4) \times (-404) + (x+4)(x+1) \times 94 + (x+4)(x+1)(x-0) \times (-14) + (x+4)(x+1)(x)(x-2) \times$$

$$= 1245 - 404x - 1616 + 94x^2 + 470x + 376 - 14x^3 - 70x^2 - 56x + 3x^4 + 9x^3 - 18x^2 - 24x$$

$$= 3x^4 - 5x^3 + 6x^2 - 14x + 5$$

2) Using Newton's divided difference formula, find $f(8)$ & $f(15)$ from the following data:

$x:$	4	5	7	10	11	13

$f(x):$	48	100	294	900	1210	2028

The divided difference table is

x	y	1^{st} divided differences	2^{nd} divided differences	3^{rd} divided differences	4^{th} divided differences
4	**48**				
		52			
5	100		**15**		
		97		**1**	
7	294		21		**0**
		202		1	
10	900		27		**0**
		310		1	
11	1210		33		
		409			
13	2028				

Newton's divided differences interpolation formula is

$$y = y_o + (x - x_o)[x_o, x_1] + (x - x_o)(x - x_1)[x_o, x_1, x_2] + (x - x_o)(x - x_1)(x - x_2)[x_o, x_1, x_2, x_3]$$

$$+ (x - x_o)(x - x_1)(x - x_2)(x - x_3)[x_o, x_1, x_2, x_3, x_4].$$

Here $[x_o, x_1] = 52,\ [x_o, x_1, x_2] = 15,\ [x_o, x_1, x_2, x_3] = 1,\ [x_o, x_1, x_2, x_3, x_4] = 0,$

$$\therefore\ f(x) = 48 + (x-4)\times 52 + (x-4)(x-5)\times 15 + (x-4)(x-5)(x-7)\times 1 + (x-4)(x-5)(x-7)(x-10)\times 0$$

$$\Rightarrow\ f(8) = 48 + (8-4)\times 52 + (8-4)(8-5)\times 15 + (8-4)(8-5)(8-7)\times 1 + (8-4)(8-5)(8-7)(8-10)\times 0$$

$$= 48 + 4\times 52 + 4\times 3\times 15 + 4\times 3\times 1\times 1 + 0 = 448\ \&$$

$$f(15) = 48 + (15-4)\times 52 + (15-4)(15-5)\times 15 + (15-4)(15-5)(15-7)\times 1 + (15-4)(15-5)(15-7)(15-10)$$

$$= 48 + 11\times 52 + 11\times 10\times 15 + 11\times 10\times 8\times 1 + 0 = 3150$$

3) Given that $f(0) = 8,\ f(1) = 68,\ \&\ f(5) = 123$. Construct a divided difference table. Using the table determine the value of $f(2)$.

Answer: The divided difference table is

x	y	1^{st} divided differences	2^{nd} divided differences
0	**8**		
		60	
1	68		**-9.25**
		13.75	
5	123		

Newton's divided differences interpolation formula is

$$y = y_o + (x - x_o)[x_o, x_1] + (x - x_o)(x - x_1)[x_o, x_1, x_2]$$

Here $[x_o, x_1] = 60,\ [x_o, x_1, x_2] = -9.25,$

$\therefore\ f(x) = 8 + (x-0)\times 60 + (x-0)(x-1)\times(-9.25)$

$\Rightarrow\ f(2) = 8 + (2-0)\times 60 + (2-0)(2-1)\times(-9.25) = 8 + 2\times 60 + 2\times 1\times(-9.25) = 109.50$

Exercises: 1) Using Newton's formula for divided differences find an interpolating polynomial for the following data:

$x:$	0.0	0.5	1.0	2.0
$f(x):$	0.00	0.57	1.46	5.05

Hence find $f(0.3)$ & $f(1.6)$.

2) Given the data:

$x:$	5	7	11	13	17
$f(x):$	150	392	1452	2366	5202

Find $f(9)$ using the Newton's divided difference formula.

3) For the data given in the following table, find the polynomial approximation of $f(x)$ using Newton's divided difference formula:

$x:$	2	4	5	6	8	10
$f(x):$	10	96	196	350	868	1746

4) Given $u_{20} = 24.37,\ u_{22} = 49.28,\ u_{29} = 162.86 \,\&\, u_{32} = 240.5$, find u_{28} by using the Newton's

divided difference formula.

5) Given $\log_{10} 300 = 2.4771,\ \log_{10} 304 = 2.4829,\ \log_{10} 305 = 2.4843, \&\log_{10} 307 = 2.4871,$ find

by divided difference formula the value of $\log_{10} 306.$

6) Given $f(0) = -18,\ f(1) = 0,\ f(3) = 0,\ f(5) = -248,\ f(6) = 0,\ f(9) = 13104,$ find $f(x).$

7) Applying the method of divided differences for interpolation, find the value of y when $x = 5,$

given

$x:$	4.50	4.55	4.70	4.90	5.15
$y:$	1345	1470	2010	3815	10965

8) If $y(1) = -3,\ y(3) = 9,\ y(4) = 30, \&\ y(6) = 132,$ find the interpolating polynomial that takes these values.

Answers: 1) $0.2834x^3 + 0.2149x^2 + 0.9617x.\ f(0.3) = 0.3155 \,\&\, f(1.6) = 3.2497$ 2) 810

3) $2x^3 - 3x^2 + 5x - 4$ 4) 141.938 5) 2.4857 6)
$x^5 - 9x^4 + 18x^3 - x^2 + 9x - 18$

7) 5745 8) $x^3 - 3x^2 + 5x - 6$

Inverse interpolation:

The process of estimating the value of x for a value of y is called the inverse interpolation. Just on interchanging x & y in the Lagrange's formula or Newton's general formula we obtain Inverse interpolation formula as:

1) $$x = \frac{(y-y_1)(y-y_2)\ldots\ldots\ldots(y-y_n)}{(y_o-y_1)(y_o-y_2)\ldots\ldots.(y_o-y_n)} \times x_o + \frac{(y-y_o)(y-y_2)\ldots\ldots\ldots(y-y_n)}{(y_1-y_o)(y_1-y_2)\ldots\ldots.(y_1-y_n)} \times x_1 + \ldots\ldots. + \frac{(y-y_0)(y-y_1)\ldots\ldots\ldots(y-y_{n-1})}{(y_n-y_o)(y_n-y_1)\ldots\ldots.(y_n-y_{n-1})} \times x_n$$

2)
$$x = x_o + (y-y_o)[y_o, y_1] + (y-y_o)(y-y_1)[y_o, y_1, y_2] + (y-y_o)(y-y_1)(y-y_2)[y_o, y_1, y_2, y_3] + \ldots\ldots\ldots\ \ldots\ldots\ldots + (y-y_o)(y-y_1)(y-y_2)\ldots..(y-y_n)[y, y_o, y_1, \ldots\ldots., y_n]$$

Problems: 1) Given

$x:$	2	5	9	11
$y:$	10	12	15	19

Find x corresponding to $y = 16$

Answer: Here $y_o = 10,\ y_1 = 12,\ y_2 = 15,\ y_3 = 19,\ x_o = 2,\ x_1 = 5,\ x_2 = 9,\ x_3 = 11,\ y = 16.$

The inverse interpolation formula is

$$x = \frac{(y-y_1)(y-y_2)(y-y_3)}{(y_o-y_1)(y_o-y_2)(y_o-y_3)} \times x_o + \frac{(y-y_o)(y-y_2)(y-y_3)}{(y_1-y_o)(y_1-y_2)(y_1-y_3)} \times x_1 + \frac{(y-y_o)(y-y_1)(y-y_3)}{(y_2-y_o)(y_2-y_1)(y_2-y_3)} \times x_2 + \frac{(y-y_o)(y-y_1)(y-y_2)}{(y_3-y_o)(y_3-y_1)(y_3-y_2)} \times x_3$$

$$= \frac{(16-12)(16-15)(16-19)}{(10-12)(10-15)(10-19)} \times 2 + \frac{(16-10)(16-15)(16-19)}{(12-10)(12-15)(12-19)} \times 5 + \frac{(16-10)(16-12)(16-19)}{(15-10)(15-12)(15-19)} \times 9 + \frac{(16-10)(16-12)(16-15)}{(19-10)(19-12)(19-15)} \times 11$$

$$= \frac{4 \times 1 \times (-3)}{(-2)(-5)(-9)} \times 2 + \frac{6 \times 1(-3)}{2(-3)(-7)} \times 5 + \frac{6 \times 4(-3)}{5 \times 3(-4)} \times 9 + \frac{6 \times 4 \times 1}{9 \times 7 \times 4} \times 11 = 9.971428577 \approx 9.97143$$

2) If $y_1 = 4,\ y_3 = 12,\ y_4 = 19, \&\ y_x = 7$ find x.

Answer: Here $y_o = 4,\ y_1 = 12,\ y_2 = 19,\ x_o = 1,\ x_1 = 3,\ x_2 = 4, \&\ y = 7.\ x = ?$

The inverse interpolation formula is

$$x = \frac{(y - y_1)(y - y_2)}{(y_o - y_1)(y_o - y_2)} \times x_o + \frac{(y - y_o)(y - y_3)}{(y_1 - y_o)(y_1 - y_3)} \times x_1 + \frac{(y - y_o)(y - y_1)}{(y_2 - y_o)(y_2 - y_1)} \times x_2$$

$$= \frac{(7-12)(7-19)}{(4-12)(4-19)} \times 1 + \frac{(7-4)(7-19)}{(12-4)(12-19)} \times 3 + \frac{(7-4)(7-12)}{(19-4)(19-12)} \times 4$$

$$= \frac{(-5)(-12)}{(-8)(-15)} \times 1 + \frac{3 \times (-12)}{8 \times (-7)} \times 3 + \frac{3 \times (-5)}{15 \times 7} \times 4 = 1.857142855 \approx 1.8571.$$

3) Apply Lagrange's formula to find a root of the equation $f(x) = 0$, given that $f(30) = -30$, $f(34) = -13$, $f(38) = 3$, $f(42) = 18$.

Answer: Here $y_o = -30,\ y_1 = -13,\ y_2 = 3,\ y_3 = 18,\ x_o = 30,\ x_1 = 34,\ x_2 = 38,\ x_3 = 42\ \&$ $y = 0.\ x = ?$

The inverse interpolation formula is

$$x = \frac{(y - y_1)(y - y_2)(y - y_3)}{(y_o - y_1)(y_o - y_2)(y_o - y_3)} \times x_o + \frac{(y - y_o)(y - y_2)(y - y_3)}{(y_1 - y_o)(y_1 - y_2)(y_1 - y_3)} \times x_1 + \frac{(y - y_o)(y - y_1)(y - y_3)}{(y_2 - y_o)(y_2 - y_1)(y_2 - y_3)} \times x_2$$

$$+ \frac{(y - y_o)(y - y_1)(y - y_2)}{(y_3 - y_o)(y_3 - y_1)(y_3 - y_2)} \times x_3$$

$$= \frac{(0+13)(0-3)(0-18)}{(-30+13)(-30-3)(-30-18)} \times 30 + \frac{(0+30)(0-3)(0-18)}{(-13+30)(-13-3)(-13-18)} \times 34 + \frac{(0+30)(0+13)(0-18)}{(3+30)(3+13)(3-18)} \times 38$$

$$+ \frac{(0+30)(0+13)(0-3)}{(18+30)(18+13)(18-3)} \times 42$$

$$= \frac{13 \times (-3)(-18)}{(-17)(-33)(-48)} \times 30 + \frac{30 \times (-3)(-18)}{17 \times (-16)(-31)} \times 34 + \frac{30 \times 13(-18)}{33 \times 16 \times (-15)} \times 38 +$$

$$\frac{30 \times 13(-3)}{48 \times 31 \times 15} \times 42 = 37.23.$$

Exercises: 1) Apply Lagrange's method to find the value of x when $f(x) = 15$ from the given data:

$x:$	5	6	9	11
$f(x):$	12	13	14	16

2) Obtain the value of t when $A = 85$ from the following table using Lagrange's method:

t	2	5	8	14
A	94.8	87.9	81.3	68.7

3) The following table given a set of values of x & y

x	1.2	2.1	2.8	4.1	4.9	6.2
y	4.2	6.8	9.8	13.4	15.5	19.6

Find the value of x corresponding to $y=12$

4) The following table gives the values of $f(\theta)$ for certain values of θ. Find θ when $f(\theta)=0.3887$

θ	21^o	23^o	25^o
$f(\theta)$	0.3706	0.4068	0.4433

5) Solve the equation $f(x)=0$ given

x	0	0.1	0.2	0.3	0.4
$f(x)$	1	0.80484	0.61873	0.44082	0.27032

Answers: 1) 11.5 2) 6.5928 3) 3.55 4) 22^o 5) 0.5671

Numerical Integration

Here for a given set of data $(x_0,y_0),(x_1,y_1),........,(x_n,y_n)$ of a function $y=f(x)$ it is required to obtain a formula to compute the value of $I=\int_a^b f(x)\ dx$ or $I=\int_a^b y\ dx$

Let $[a,b]$ be divided into n equal subintervals such that $a=x_o<x_1<x_2<..........<x_n=b$.

Then clearly $x_n=x_o+nh$. Then integral becomes $I=\int_a^b f(x)\ dx=\int_{x_o}^{x_n} f(x)\ dx$

By considering Newton's forward difference formula

$$f(x)=y_0+\frac{p\Delta y_0}{1!}+\frac{p(p-1)\Delta^2 y_0}{2!}+\frac{p(p-1)(p-2)\Delta^3 y_0}{3!}+................$$

$$I=\int_{x_o}^{x_n}\left[y_0+\frac{p\Delta y_0}{1!}+\frac{p(p-1)\Delta^2 y_0}{2!}+\frac{p(p-1)(p-2)\Delta^3 y_0}{3!}+...............\right]dx$$

Since $x=x_o+ph,$ we have $dx=hdp$, then if $x=x_o\Rightarrow p=0\ \&\ if\ x=x_n\Rightarrow p=n.$

$$\therefore I=h\int_0^n\left[y_0+p\Delta y_o+\frac{p(p-1)\Delta^2 y_0}{2}+\frac{p(p-1)(p-2)\Delta^3 y_0}{6}+...............\right]dp$$

$$=h\left[y_o p+\frac{p^2}{2}\Delta y_o+\frac{1}{2}\left(\frac{p^3}{3}-\frac{p^2}{2}\right)\Delta^2 y_o+\frac{1}{6}\left(\frac{p^4}{4}-\frac{3p^3}{3}+\frac{2p^2}{2}\right)\Delta^3 y_0+..........\right]_0^n$$

$$= nh\left[y_o + \frac{n}{2}\Delta y_o + \frac{1}{12}n(2n-3)\Delta^2 y_o + \frac{1}{24}n(n-2)^2\Delta^3 y_0 + \dots\dots\dots\right] \text{----------(1)}$$

This formula which gives the value of I is known as the **quadrature formula**. By putting $n = 1,\ n = 2,\ n = 3,$ …….. we get different integration formulas.

1) **Simpson's Rule**: By putting $n = 2$ in Equation (1) we get,

$$\int_{x_o}^{x_2} y\,dx = 2h\left[y_o + \frac{2}{2}\Delta y_o + \frac{1}{12}2(4-3)\Delta^2 y_o\right] = 2h\left[y_o + \Delta y_o + \frac{1}{6}\Delta^2 y_0\right]$$

$$= 2h\left[y_o + y_1 - y_o + \frac{1}{6}(y_2 - 2y_1 + y_o)\right] = \frac{h}{3}[y_o + 4y_1 + y_2]$$

Similarly $\int_{x_2}^{x_4} y\,dx = \frac{h}{3}[y_2 + 4y_3 + y_4]$ & soon finally $\int_{x_{n-2}}^{x_n} y\,dx = \frac{h}{3}[y_{n-2} + 4y_{n-1} + y_n]$

$\therefore$ By adding we get

$$\int_{x_0}^{x_n} y\,dx = \frac{h}{3}[(y_o + y_n) + 4(y_1 + y_3 + \dots\dots + y_{n-1}) + 2(y_2 + y_4 + \dots\dots + y_{n-2})]]$$

This is called **Simpson's 1/3 – rule or Simpson's rule**.

Note: This rule is application only when n is an even number.

2) **Simpson's Three- eight Rule**: By putting $n = 3$ in Equation (1) we get,

$$\int_{x_o}^{x_3} y\,dx = 3h\left[y_o + \frac{3}{2}\Delta y_o + \frac{3}{4}\Delta^2 y_o + \frac{1}{8}\Delta^3 y_0\right]$$

$$= \frac{3h}{8}[y_o + 3y_1 + 3y_2 + y_3]$$

Similarly we get $\int_{x_3}^{x_6} y\,dx = \frac{3h}{8}[y_3 + 3y_4 + 3y_5 + y_6]$, & soon. Finally we get

$\int_{x_{n-3}}^{x_n} y\,dx = \frac{3h}{8}[y_{n-3} + 3y_{n-2} + 3y_{n-1} + y_n]$, By adding we get,

$$\int_{x_o}^{x_n} y\,dx = \frac{3h}{8}[(y_o + y_n) + 3(y_1 + y_2 + y_4 + y_5 + y_7 + \dots\dots + y_{n-2} + y_{n-1}) + 2(y_3 + y_6 + \dots\dots + y_{n-3})]$$

This is called as **Simpson's** $(3/8)^{th}$ **rule** & is applicable only when n is a multiple of 3.

Weddle's Rule: By putting $n = 6$ in Equation (1) we get Weddle's formula as

$$\int_{x_o}^{x_n} y\,dx = \frac{3h}{10}[(y_o + y_n) + 5(y_1 + y_5 + \dots. + y_{n-5} + y_{n-1}) + (y_2 + y_4 + \dots.. + y_{n-4} + y_{n-2})$$

$$+ 2(y_6 + y_{12} + \dots.. + y_{n-6}) + 6(y_3 + y_9 + \dots.. + y_{n-3})]$$

This is called as **Weddle's rule** & is generally used when the interval $[x_o, x_n]$ is divided into **6 or 12** equal parts.

Weddle's rule for n = 6 is given by

$$\int_{x_o}^{x_6} y\,dx = \frac{3h}{10}\left[(y_o + y_6) + 5(y_1 + y_5) + (y_2 + y_4) + 6y_3\right]$$ & for n = 12 is given by

$$\int_{x_o}^{x_{12}} y\,dx = \frac{3h}{10}\left[(y_o + y_6) + 5(y_1 + y_5 + y_7 + y_{11}) + (y_2 + y_4 + y_8 + y_{10}) + 2y_6 + 6(y_3 + y_9)\right]$$

Problems: 1) By using Simpson's 1/3 rule, evaluate $\int_0^1 \frac{dx}{1+x^2}$, dividing the interval [0, 1] into six equal

parts. Hence find an approximate value of π .

- Here the interval [0, 1] is to be divided into 6 equal parts. Hence n = 6.

$$\Rightarrow h = \frac{b-a}{n} = \frac{1-0}{6} = \frac{1}{6}$$

& $y = \frac{1}{1+x^2}$. Hence the functional values at $x_0 = 0, x_1 = \frac{1}{6}, x_2 = \frac{1}{3}, \ldots\ldots\ldots, x_6 = 1$ are shown below

x:	$x_0 = 0$	$x_1 = 1/6$	$x_2 = 1/3$	$x_3 = 1/2$	$x_4 = 2/3$	$x_5 = 5/6$	$x_6 = 1$
f(x) :	$y_0 = 1$	$y_1 = 0.97297$	$y_2 = 0.9$	$y_3 = 0.8$	$y_4 = 0.69231$	$y_5 = 0.59016$	$y_6 = 0.5$

The Simpson's 1/3 rule is

$$\int_0^1 \frac{dx}{1+x^2} = \frac{h}{3}\left[(y_0 + y_6) + 4(y_1 + y_3 + y_5) + 2(y_2 + y_4)\right]$$

$$= \frac{1/6}{3}\left[(1+0.5) + 4(0.97297 + 0.8 + 0.59016) + 2(0.9 + 0.69231)\right] = 0.785397$$

The exact value of the integral is $\int_0^1 \frac{dx}{1+x^2} = \left[\tan^{-1} x\right]_0^1 = \tan^{-1} 1 - \tan^{-1} 0 = \frac{\pi}{4}$

Hence $\frac{\pi}{4} = 0.785397 \Rightarrow \pi = 0.785397 \times 4 = \mathbf{3.141588}$

2) Evaluate $\int_0^{\pi/2} \sqrt{\cos\theta}\, d\theta$ by dividing the interval into eight equal parts.

- Here n = 8 $\Rightarrow h = \frac{\pi/2 - 0}{8} = \frac{\pi}{16}$. $f(\theta) = \sqrt{\cos\theta}$ $\therefore$ the table of functional values is

  given by

θ:	0	$\pi/16$	$\pi/8$	$3\pi/16$	$\pi/4$	$5\pi/16$	$3\pi/8$	$7\pi/16$	$\pi/2$

y = $f(\theta)$:	1	0.9903	0.9612	0.9118	0.8409	0.7454	0.6186	0.4417	0

The Simpson's 1/3 rule is

$$\int_0^{\pi/2}\sqrt{\cos\theta}\,d\theta=\frac{h}{3}\left[(y_0+y_8)+4(y_1+y_3+y_5+y_7)+2(y_2+y_4+y_6)\right]$$

$$=\frac{\pi/16}{3}\left[(1+0)+4(0.9903+0.9118+0.7454+0.4417)+2(0.9612+0.8409+0.6186)\right]$$
$$=\mathbf{1.19155}$$

3) By using Simpson's 1/3 rule, evaluate $\int_0^1 e^{-x^2}dx$, taking h = 0.1

- Here $f(x)=e^{-x^2}$. Hence the functional values are-

x:	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
f(x):	1	0.9901	0.9608	0.9139	0.8521	0.7788	0.6977	0.6126	0.5273	0.4449	0.3679

The Simpson's 1/3 rule is

$$\int_0^1 e^{-x^2}dx=\frac{h}{3}\left[(y_0+y10)+4(y_1+y_3+y_5+y_7+y_9)+2(y_2+y_4+y_6+y_8)\right]$$

$$=\frac{0.1}{3}\begin{bmatrix}(1+0.3679)+4(0.9901+0.9139+0.7788+0.6126+0.4449)\\+2(0.9608+0.8521+0.6977+0.5273)\end{bmatrix}=\mathbf{0.74683}$$

4) A solid of revolution is formed by rotating about the x – axis the area between the x –axis, the lines x = 0

and x = 1, and a curve through the points with the following coordinates.

x:	0.00	0.25	0.50	0.75	1.00
y:	1.0000	0.9896	0.9589	0.9089	0.8415

Using the Simpson's rule, find the volume of the solid formed.

- The required volume is given by $V=\pi\int_0^1 y^2dx$

First we have to find the values of $Y=y^2$ for the specified values of x which is as shown below-

x:	0.00	0.25	0.50	0.75	1.00
$Y = y^2$:	1.0000	0.9793	0.9195	0.8261	0.7081

The Simpson's 1/3 rule is

$$V = \pi \times \frac{h}{3}\left[(Y_0 + Y_4) + 4(Y_1 + Y_3) + 2Y_2\right]$$

$$= \frac{\pi \times 0.25}{3}\left[(1 + 0.7081) + 4(0.9793 + 0.8261) + 2 \times 0.9195\right] = \mathbf{2.8192}$$

5) Evaluate $\int_0^1 \frac{dx}{1+x}$ by applying the Simpson's 3/8 rule with seven ordinates. Hence, deduce the

value of $\log 2$.

- Note that taking seven ordinates is equivalent to dividing the interval into six equal parts.

n = 6 $\Rightarrow h = \frac{1-0}{6} = \frac{1}{6}$ & $f(x) = \frac{1}{x+1}$, the functional values are as shown below-

x:	$x_0 = 0$	$x_1 = 1/6$	$x_2 = 1/3$	$x_3 = 1/2$	$x_4 = 2/3$	$x_5 = 5/6$	$x_6 = 1$
f(x) :	$y_0 = 1$	$y_1 = 0.8571$	$y_2 = 0.75$	$y_3 = 0.6667$	$y_4 = 0.6$	$y_5 = 0.5455$	$y_6 = 0.5$

Now Simpson's 3/8 rule is

$$\int_0^1 \frac{dx}{1+x} = \frac{3h}{8}\left[(y_0 + y_6) + 3(y_1 + y_2 + y_4 + y_5) + 2y_3\right]$$

$$= \frac{3 \times 1/6}{8}\left[(1 + 0.5) + 3(0.8571 + 0.75 + 0.6 + 0.5455) + 2 \times 0.6667\right] = \mathbf{0.6932}$$

By direct integration we get $\int_0^1 \frac{dx}{1+x} = \log(1+x)\Big|_0^1 = \log 2 - \log 1 = \log 2$

$\therefore \log 2 = \mathbf{0.6932}$

6) By using Simpson's 3/8 rule, evaluate $\int_0^{\pi/2} e^{\sin\theta} d\theta$

- Let n = 6, then $h = \frac{\pi}{12}$. $f(\theta) = e^{\sin\theta}$. $\therefore$ the functional values are-

θ :	0	$\pi/12$	$\pi/6$	$\pi/4$	$\pi/3$	$5\pi/12$	$\pi/2$
$f(\theta) = e^{\sin\theta}$:	1	1.2954	1.6487	2.0281	2.3774	2.6272	2.7183

Now Simpson's 3/8 rule is

$$\int_0^{\pi/2} e^{\sin\theta}\,d\theta = \frac{3h}{8}\left[(y_0+y_6)+3(y_1+y_2+y_4+y_5)+2y_3\right]$$

$$=\frac{3}{8}\times\frac{\pi}{12}\left[(1+2.7183)+3(1.2954+1.6487+2.3774+2.6272)+2\times 2.0281\right]=\mathbf{3.1043}$$

7) By using Simpson’s 3/8 rule with $h=0.2$ find the approximate area under the curve

$y=\frac{x^2-1}{x^2+1}$ between the

ordinates $x=1\;\&\;x=2.8.$ Compare the result with the exact result.

- Here required area under the curve is $A=\int_1^{2.8}\frac{x^2-1}{x^2+1}dx$. So $f(x)=\frac{x^2-1}{x^2+1}$. The functional values are

$x:$	1	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.6	2.8
$f(x):$	0	0.1803	0.3243	0.4382	0.5283	0.6	0.6575	0.7041	0.7423	0.7738

Now Simpson’s 3/8 rule is

$$A=\int_1^{2.8}\frac{x^2-1}{x^2+1}dx=\frac{3h}{8}\left[(y_0+y_9)+3(y_1+y_2+y_4+y_5+y_7+y_8)+2(y_3+y_6)\right]$$

$$=\frac{3\times 0.2}{8}\left[0+0.7738+3(0.1803+0.3243+0.5283+0.6+0.7041+0.7423)+2(0.4382+0.6575)\right]$$

$$=\mathbf{0.9152}$$

Exact value is $A=\int_1^{2.8}\frac{x^2-1}{x^2+1}dx=\int_1^{2.8}\frac{x^2+1-2}{x^2+1}dx=\int_1^{2.8}\left(1-\frac{2}{x^2+1}\right)dx=\left[x-2\tan^{-1}x\right]_1^{2.8}$

$$=2.8-2\tan^{-1}2.8-\left(1-2\tan^{-1}1\right)=\mathbf{0.9152}$$

8) Given

$x:$	4.0	4.2	4.4	4.6	4.8	5.0	5.2
$\log x:$	1.3863	1.4351	1.4816	1.5261	1.5286	1.6094	1.6487

Evaluate $\int_4^{5.2}\log x\;dx$, by using the Weddle’s rule.

- Here n = 6 & h = 0.2. So we can use Weddle’s rule-

The rule is $\int_4^{5.2} \log x dx = \frac{3h}{10}\left[(y_o + y_6) + 5(y_1 + y_5) + (y_2 + y_4) + 6y_3\right]$

$$= \frac{3\times 0.2}{10}\left[(1.3863 + 1.6487) + 5(1.4351 + 1.6094) + (1.4816 + 1.5286) + 6\times 1.5261\right]$$

=1.8255

9) Evaluate $\int_0^{\pi/2} \sqrt{1-0.162\sin^2\theta}\ d\theta$ by using Weddle's rule with 7 ordinates.

- Here 7 ordinates means n = 6 & so $h = \frac{\pi}{12}$ & here $f(\theta) = \sqrt{1-0.162\sin^2\theta}$. So the functional values are –

θ	0	$\pi/12$	$\pi/6$	$\pi/4$	$\pi/3$	$5\pi/12$	$\pi/2$
$f(\theta)$	1	0.9946	0.9795	0.9586	0.9373	0.9213	0.9154

The Weddle's rule is

$$\int_0^{\pi/2} \sqrt{1-0.162\sin^2\theta} d\theta = \frac{3h}{10}\left[(y_o + y_6) + 5(y_1 + y_5) + (y_2 + y_4) + 6y_3\right]$$

$$= \frac{3\times\pi}{10\times 12}\left[(1 + 0.9154) + 5(0.9946 + 0.9213) + (0.9795 + 0.9373) + 6\times 0.9586\right]$$

$= \mathbf{1.5052}$

10) By using the Weddle's rule find the total arc length (perimeter) of the ellipse $x^2 + y^2/4 = 1$

- The parametric equation of the ellipse is $x = \cos\theta \;\&\; y = 2\sin\theta$

The length of the curve is given by $L = 4\int_0^{\pi/2} \sqrt{\left(\frac{dx}{d\theta}\right)^2 + \left(\frac{dy}{d\theta}\right)^2}\, d\theta$

Here $\frac{dx}{d\theta} = -\sin\theta \;\&\; \frac{dy}{d\theta} = 2\cos\theta \Rightarrow$

$$\sqrt{\left(\frac{dx}{d\theta}\right)^2 + \left(\frac{dy}{d\theta}\right)^2} = \sqrt{(-\sin\theta)^2 + (2\cos\theta)^2} = \sqrt{\sin^2\theta + 4\cos^2\theta}$$

$$= \sqrt{1+3\cos^2\theta}$$

Hence $L = 4\int_0^{\pi/2} \sqrt{1+3\cos^2\theta}\ d\theta$

Now dividing the interval into 6 equal parts we get $h = \frac{\pi}{12}$ & $f(\theta) = \sqrt{1+3\cos^2\theta}$. So the functional values are –

θ	0	$\pi/12$	$\pi/6$	$\pi/4$	$\pi/3$	$5\pi/12$	$\pi/2$
$f(\theta)$	2	1.9491	1.8028	1.5811	1.3229	1.0959	1

Hence by Weddle's rule we get,

$$L = 4\int_0^{\pi/2} \sqrt{1+3\cos^2\theta}\,d\theta = \frac{12h}{10}\left[(y_o+y_6)+5(y_1+y_5)+(y_2+y_4)+6y_3\right]$$

$$= \frac{\pi}{10}\left[(2+1)+5(1.9491+1.0959)+(1.8028+1.3229)+6\times 1.5811\right] = \mathbf{9.6878}$$

Exercises:

1. By using the Simpson's 1/3 rule, evaluate $\int_1^5 f(x)dx$, given

x	1	2	3	4	5
f(x)	13	50	70	80	100

1. Evaluate $\int_0^1 \frac{xdx}{1+x^2}$ by using Simpson's 1/3 rule taking six equal strips, & hence deduce an approximate value of log 2.
2. Evaluate $\int_0^5 \frac{dx}{4x+5}$ by using the Simpson's 1/3 rule with the aid of 11 ordinates. Hence find log 5.

4. A rocket is launched from the ground. Its acceleration f is registered during the first 80 seconds and is tabulated below. Using the Simpson's rule, find the velocity of the rocket at t= 80 seconds.

t (sec)	0	10	20	30	40	50	60	70	80
f (cm/sec^2)	30	31.63	33.34	35.47	37.75	40.33	43.25	46.69	50.67

5. Use Simpson's rule with h = 0.125 to show that $\int_0^1 e^{-x^2/2}\,dx = 0.85581$.

6. Evaluate $\int_0^6 \frac{dx}{1+x^2}$ by using the Simpson's 3/8 rule.

7. Using Simpson's 3/8 rule evaluate $\int_0^{0.3} \sqrt{1-8x^3}\,dx$, taking 7 ordinates.

8. Evaluate $\int_0^1 \frac{x}{1+x^2}dx$ by using the Simpson's 3/8 rule, dividing the interval into 3 equal parts.
Hence find an approximate value of $\log\sqrt{2}$.

9. Evaluate $\int_0^{\pi/2} \sqrt{\sin x}\ dx$ using the Simpson's 3/8 rule, with 3 subintervals of $\left(0, \pi/2\right]$.

10. Evaluate $\int_0^3 \frac{dx}{(1+x)^2}$ by using Simpson's 3/8 rule with six subintervals of [0, 3].

11. Evaluate the integral $\int_0^6 \frac{dx}{1+x^2}$ by using the Weddle's rule with h = 0.5. Compare the result with the actual value.

12. Evaluate $\int_0^1 \frac{dx}{1+x^2}$ by using Weddle's rule with seven ordinates. Hence find an approximate value of π.

13. By using Weddle's rule evaluate $\int_{0.2}^{1.4} (\sin x - \log x + e^x)dx.$ Divide the interval into 12 equal parts.

14. Using Weddle's rule, find the area under the curve $y = f(x)$ between the ordinates x = 0 & x = 6, the function $f(x)$ being described by the following table:

x	0	1	2	3	4	5	6
f(x)	0	2	2.5	2.3	2	1.7	1.5

15. By using Weddle's rule, evaluate $\int_{0.4}^{1.6} x\cos cehx\ dx$, dividing the interval into 12 equal parts.

Answers: 1) 257.6667 2) 0.34661, 0.69

◈◈◈◈◈

Numerical Methods-II

Lesson Structure

Numerical solution of ordinary differential equations of first order and first degree:

1.1 Introduction

1.2 Picard’s Method of successive approximation:

1.3 Taylor’s Series Method

1.4 Modified Euler’s Method

1.5 Runge - Kutta Method

1.6 Milne’s Method

1.7 Course Outcome

1.8 Exercise

1.9 Further Reading

1.1 Introduction

A ***numerical method*** can be used to get an accurate approximate solution to a differential equation. There are many programs and packages available for solving these differential equations. With today's computer, an accurate solution can be obtained rapidly. In this chapter we focus on basic numerical methods for solving initial value problems.

Analytical methods, when available, generally enable to find the value of y for all values of x. Numerical methods, on the other hand, lead to the values of y corresponding only to some finite set of values of x. That is the solution is obtained as a table of values, rather than as continuous function. Moreover, analytical solution, if it can be found, is exact, whereas a numerical solution inevitably involves an error which should be small but may, if it is not controlled, swamp the true solution. Therefore, we must be concerned with two aspects of numerical solutions of ODEs: both the method itself and its accuracy.

The most general form of an ODE of first order and first degree is given by

$$\frac{dy}{dx} = f(x,y) where\ y(x_0) = y_0$$

Let x by an independent variable and y be dependent variable.

Let us consider the differential equation $\frac{dy}{dx} = f(x,y) where\ y(x_0) = y_0$-------(1)

If particular values are given to the constants, then the resulting solution is called a particular solution.

To obtain a particular solution from the general solution (1), we must be given initial conditions so that the constants can be determined. If all the initial conditions are specified at the same value of x then the problem is termed as initial value problem. If the conditions are specified at more than one value of x, then the problem is termed as boundary value problem.

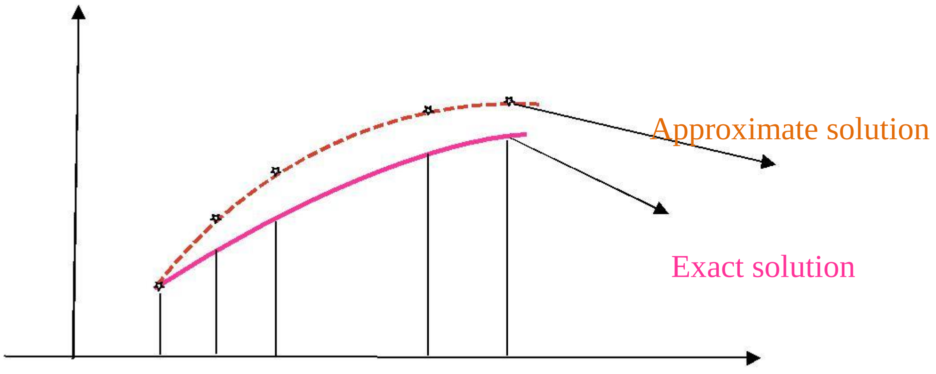
Though there are many analytical methods for finding the solution of the equation of the form (1), there exist large number of ODE's whose solution cannot be obtained by the known analytical methods. In such cases, we use numerical methods to get an approximate solution of a given differential equation under the prescribed conditions

Consider the first order differential equation $\frac{dy}{dx} = f(x, y)$

Let $y(x_0)$, $y(x_1)$, $y(x_2)$, $y(x_3)$..... $y(x_m)$ be the solution values at the points x_0, x_1, x_2, x_3,----- x_m

We wish to find the approximate values y_0 , y_1, ..., y_m to these solution values.

Let the initial condition be $y(x_0) = y_0$. Let the exact solution y(x) of the given differential equation be represented by a continuous curve. Divide the interval (Xo,Xm))on which the solution is derived into a finite number of equispaced subintervals.



For each x_i , the approximate values of the dependent variable y(x) are calculated using a suitable recursive formula. These values are y_0 , y_1, ..., y_m

and these are shown by points. Computation of these approximate values is known as Numerical solution of the Differential equation.

The general form of first order differential equation, in implicit form, is $F(x, y, y') = 0$ and in the explicit form is $\frac{dy}{dx} = f(x, y)$. An **Initial Value Problem** (IVP) consists of a differential equation and a condition which the solution much satisfies (or several conditions referring to the same value of x if the differential equation is of higher order).

$\frac{dy}{dx} = f(x,y) where\ y(x_0) = y_0$-------(1)

Assuming f to be such that the problem has a unique solution in some interval containing x_0, we shall discuss the methods for computing numerical values of the solution. These methods are **step-by-step methods**. That is, we start = from y_0 and proceed stepwise. In the first step, we compute an approximate value y_1 of the solution y of (1) at $x = x_1 = x_0 + h$. In the second step we compute an approximate value y_2 of the solution y at $x = x_2 = x_0 + 2h$, etc. Here h is fixed number for example 0.1or 0.5 depends on the requirement of the problem. In each step the computations are done by the same formula.

The following methods are used to solve the IVP (1).

1. Picard's Method
2. Taylor's Series Method
3. Euler and Modified Euler Method
4. Runge – Kutta Method
5. Milne's Method
6. Adams – Bashforth Method

1.2 Picard's Method of successive approximation:

Consider a differential equation $\frac{dy}{dx} = f(x)$ with the initial condition $y(x_0)=y_0$. The formula to find the successive approximation by picard's is given by

$$y_1 = y_0 + \int_{x_0}^{x} f(x, y_0)dx$$

$$y_2 = y_0 + \int_{x_0}^{x} f(x, y_1)dx$$

$$y_2 = y_0 + \int_{x_0}^{x} f(x, y_2)dx$$

And so on

Example 1:

Using picard's process of successive approximation, obtain a solution upto the fifth approximation of the equation $\frac{dy}{dx} = x + y$ such that y=1 when x=0 and hence find y(0.1).

Solution: $y_1 = y_0 + \int_{x_0}^{x} f(x, y_0)dx = 1 + \int_0^x f(x, 1)dx$

$=1+\int_0^x (x+1)dx = 1+\frac{x^2}{2}+x$

$$y_{2}{=}y_{0+}\int_{x_0}^{x} f(x,y_1)dx = 1+\int_0^x x+y_1 dx$$

$$=1+\int_0^x\left(x+1+\frac{x^2}{2}+x\right)dx = 1+\int_0^x\left(2x+1+\frac{x^2}{2}\right)dx$$

$$=x+1+x^2+\frac{x^3}{6}$$

$$y_3 = y_{0+}\int_{x_0}^{x} f(x,y_1)dx = 1+\int_0^x x+y_2 dx$$

$$=1+\int_0^x\left(x+x+1+x^2+\frac{x^3}{6}\right)dx = 1+\int_0^x\left(2x+1+x^2+\frac{x^3}{6}\right)dx$$

$= 1+x+x^2+\frac{x^3}{6}+\frac{x^4}{24}$

$$y_4 = y_{0+}\int_{x_0}^{x} f(x,y_3)dx = 1+\int_0^x x+y_3 dx$$

$= 1+\int_0^x\left(1+2x+x^2+\frac{x^3}{6}+\frac{x^4}{24}\right)dx$ Example :1$= 1+x+x^2+\frac{x^3}{6}+\frac{x^4}{12}+\frac{x^5}{120}$

$$y_4 = y_{0+}\int_{x_0}^{x} f(x,y_3)dx = 1+\int_0^x x+1+x+x^2+\frac{x^3}{6}+\frac{x^4}{12}+\frac{x^5}{120}\,dx$$

$= 1+x+x^2+\frac{x^3}{3}+\frac{x^4}{12}+\frac{x^5}{60}+\frac{x^5}{720}$

y(0.1)=1.1103

Example 2:

Using picard's process of successive approximation, to obtain correct to four decimal places, solution for the differential equation $\frac{dy}{dx}=x^2+y^2$ for x=0.4 given that y=0 when x=0

Solution:

$$y_{1}=y_0 + \int_{x_0}^{x} f(x,y_0)dx = 0 + \int_{0}^{x} (x^2 + {y_0}^2)dx = \int_{0}^{x} (x^2)dx = \frac{x^3}{3}$$

$$y_2 = y_0 + \int_{x_0}^{x} f(x,y_1)dx = 0 + \int_{0}^{x} (x^2 + \left[\frac{x^3}{3}\right]^2)dx = \int_{0}^{x} (x^2 + \frac{x^6}{9})dx = \frac{x^3}{3} + \frac{x^7}{63}$$

$$y_3 = y_0 + \int_{x_0}^{x} f(x,y_2)dx = 0 + \int_{0}^{x} \left[x^2 + \frac{x^6}{9} + \frac{x^{14}}{3969} + \frac{2x^{10}}{189}\right] dx$$

$=\frac{x^3}{3} + \frac{x^7}{63} + \frac{x^{15}}{59535} + \frac{2x^{11}}{2079}$

At x=0.4

Y(0.4)=$\frac{0.4^3}{3} + \frac{0.4^7}{63} + \frac{0.4^{15}}{59535} + \frac{2(0.4)^{11}}{2079} = 0.0214$

1.3 Taylor's Series Method

Let y=f(x) be a solution of the equation

Expanding it by Taylor's series about x - x_0 we get

$$f(x) = y_0 + \frac{(x - x_0)}{1!}{y_0}^1 + \frac{(x - x_0)^2}{2!}{y_0}^2 + \frac{(x - x_0)^3}{3!}{y_0}^3 + \cdots.$$

$$f(x_{n+1}) = y_n + \frac{h}{1!}{y_n}^1 + \frac{h^2}{2!}{y_n}^2 + \frac{h^3}{3!}{y_n}^3 + \frac{h^4}{4!}{y_n}^4 + \cdots.$$

.

Example-1 .

Compute y at x=1.1 and 1.2 using Taylor's series method to correct to 4 decimal places for $y^1 = x + y, y(o) = 0$

Solution:

$y_1 = x + y$ $\qquad y_1(0) = 1$

Differentiate with respect to x we get

$y_2 = 1 + y_1$ $\qquad y_2(0) = 2$

$y_3 = y_2$ $\qquad y_3\,(0) = 2$

$$f(x) = y_0 + \frac{(x-x_0)}{1!} y_1(0) + \frac{(x-x_0)^2}{2!} y_2(0) + \frac{(x-x_0)^3}{3!} y_3\,(0) + \cdots.$$

$y_1 = f(x_1) = y_0 + \frac{h}{1!} {y_0}^1 + \frac{h^2}{2!} {y_0}^2 + \frac{h^3}{3!} {y_0}^3 + \frac{h^4}{4!} {y_0}^4 + \cdots.$----------(2)

From the Taylor's series, we have h= $x - x_0$=1.1-1=0.1

Substituting all these values in Eqn(2) we get

$y_1 = f(1.1) = 0 + \frac{0.1}{1!}(1) + \frac{0.1^2}{2!} 2 + \frac{0.1^3}{3!}(2) + \cdots..$

$y_1 = 0.1103$

∴ y_1 =y(1.1) = 0.1103

$y_2 = f(x_2) = y_1 + \frac{h}{1!} {y_1}^1 + \frac{h^2}{2!} {y_1}^2 + \frac{h^3}{3!} {y_1}^3 + \frac{h^4}{4!} {y_1}^4 + \cdots.$-

y_1^{11}=x_1 + y_1 y^1(1.1)= 1.1+ 0.1103 + 1.2103

Differentiate with respect to x we get

y_1^{11}=1 + y_1^1 y^{11}(1.1)= 1+ 1.2103 + 2.2103

Y^{111}= Y^{11} y^{111}(0)=2.2103

$y_2 = f(1.2) = 0.1103 + \frac{0.1}{1!} 1.2103 + \frac{0.1^2}{2!} 2.2103 + \frac{0.1^3}{3!} 2.2103 + \cdots.$-=0.2427

Example-2:

Using the Taylor's series method , find an approximate solution correct to four decimals at x=0.1 for the IVP$\frac{dy}{dx} = x - y^2, y(0) = 1$

Solution:

$y_1 = x - y^2$ $y_1(0) = 0 - 1^2 = -1$

Differentiating w.r.to.x we get

$y_2 = 1 - 2yy_1$ $\quad$ $y_2(0) = 1 - 2 * 1 * -1_1 = -3$

$y_3 = -(2yy_2 + 2y_1^2)$ $\quad$ $y_3(0) = -(2 * 1 * -3 + 2 * (-1)^2 = -8$

$y_4 = -2(yy_3 + 3y_1y_2)$ $\quad$ $y_4(0) = -2(1 * -8 + 3 * -1 * -3) = 34$

$$f(x) = y_0 + \frac{(x - x_0)}{1!}{y_0}^1 + \frac{(x - x_0)^2}{2!}{y_0}^2 + \frac{(x - x_0)^3}{3!}{y_0}^3 + \cdots.$$

Where x=0.1 and $x_0 = 0$ we get

$y_1 = f(0.1) = 1 + \frac{0.1}{1!}(-1) + \frac{0.1^2}{2!}(-3) + \frac{0.1^3}{3!}(-8) + \frac{0.1^4}{4!}(34)$ .....

Y(0.1)=0.91379

Example-3:

Using the Taylor's series method , find an approximate solution correct to four decimals at x=0.1 for the IVP$\frac{dy}{dx} = x^2y - 1, y(0) = 1$

Solution:

$y_1 = x^2y - 1$ $\quad$ $y_1(0) = 0 - 1 = -1$

Differentiating w.r.to.x we get

$y_2 = 2xy + x^2y_1$ $\quad$ $y_2(0) = 2 * 0 * 1 + 0 = 0$

$y_3 = 2xy_1 + 2y + 2xy_1 + x^2y_2$ add the same term we get

$y_3 = 4xy_1 + 2y + x^2y_2$ $\quad$ $y_3(0) = 4 * 0 * -1 + 2 * 1 + 0 = 2$

$y_4 = 4xy_2 + 4y_1 + 2y_1 + x^2y_3 + 2xy_2$

$y_4 = 6xy_2 + 6y_1 + x^2y_3$ $\quad$ $y_4(0) = 0 + 6(-1) + 0 = -6$

$$f(x) = y_0 + \frac{(x - x_0)}{1!}{y_0}^1 + \frac{(x - x_0)^2}{2!}{y_0}^2 + \frac{(x - x_0)^3}{3!}{y_0}^3 + \cdots.$$

Where x=0.1 and $x_0 = 0$ we get

$$y_1 = f(0.1) = 1 + \frac{0.1}{1!}(-1) + \frac{0.1^2}{2!}(0) + \frac{0.1^3}{3!}(2) + \frac{0.1^4}{4!}(-6) + \cdots$$

Y(0.1)=0.9003

Example-4:

Using the Taylor's series method , find an approximate solution correct to four decimals at x=0.1 for the IVP$\frac{dy}{dx} = 2y + 3e^x, y(0) = 0$

Solution:

$y_1 \ = 2y + 3e^x$ $\qquad y_1(0) = 0 + 3e^0 = 3$

Differentiating w.r.to.x we get

$y_2 = 2y_1 + 3e^x$ $\qquad y_2(0) = 2 * 3 + 3e^0 = 9$

$y_3 = 2y_2 + 3e^x$ $\qquad y_3(0) = 2 * 9 + 3e^0 = 21$

$y_4 = 2y_3 + 3e^x$ $\qquad y_4(0) = 2 * 21 + 3e^0 = 45$

$$f(x) = y_0 + \frac{(x - x_0)}{1!} {y_0}^1 + \frac{(x - x_0)^2}{2!} {y_0}^2 + \frac{(x - x_0)^3}{3!} {y_0}^3 + \cdots.$$

Where x=0.1 and $x_0 = 0$ we get

$$y_1 = f(0.1) = 0 + \frac{0.1}{1!}(3) + \frac{0.1^2}{2!}(9) + \frac{0.1^3}{3!}(21) + \frac{0.1^4}{4!}(45) + \cdots$$

Y(0.1)=0.3487

1.4 Modified Euler's Method

Consider the IVP $\frac{dy}{dx} = f(x, y), y(0) = y_0$ ----------------------(1)

To determine the solution of this problem at $x_n = x_0 + nh$ bu using Euler's method.

$y_n^{(E)} = y_{n-1} + hf(x_{n-1}, y_{n-1})$--------------------------------------(2)

The expression 2 gives an approximate value of y at x_n. . To improve the approximateion the following formula has been suggested

$y_n = y_{n-1} + \frac{h}{2}\left[f(x_{n-1}, y_{n-1}) + f(x_n, y_n^{(E)})\right]$---------------------------(3)

The method of computing y_n using (3) is known as Modified Euler's method.

The process of improving the approximation can be continued by obtaining replacing $y_n^{(1)}, y_n^{(2)}, y_n^{(3)}$ until the desired degree of accuracy is obtained.

First approximation

$$y_1^{(E)} = y_0 + hf(x_0, y_0)$$

Second approximation $y_1^{(1)} = y_0 + \frac{h}{2}\left[f(x_0, y_0) + f(x_1, y_1^{(E)})\right]$

Third approximation $y_1^{(2)} = y_0 + \frac{h}{2}\left[f(x_0, y_0) + f(x_1, y_1^{(1)})\right]$ and so on

Example :1

Using the modified Euler's method, solve the IVP $\frac{dy}{dx} = \frac{1}{x+y}, y(0) = 1$ at points x=0.5 and x=1 in steps of length h=0.5. Carry out two modifications at each step.

Solution:

$\frac{dy}{dx} = f(x, y) = \frac{1}{x+y}$ $x_0 = 0\,, y_0 = 1$ $taking\ h = 0.5$

$y_1^{(E)} = y_0 + hf(x_0, y_0) = y_0 + h\frac{1}{x_0+y_0} \quad = 1 + \frac{0.5}{0+1} = 1.5$

Second approximation $y_1^{(1)} = y_0 + \frac{h}{2}\left[f(x_0, y_0) + f(x_1, y_1^{(E)})\right]$

$$= y_0 + \frac{h}{2}\left[\frac{1}{x_0 + y_0} + \frac{1}{x_0 + y_1^E}\right] = 1 + \frac{0.5}{2}\left[\frac{1}{0+1} + \frac{1}{0.5 + 1.5} = 1.375\right]$$

Third approximation $y_1^{(2)} = y_0 + \frac{h}{2}\left[f(x_0, y_0) + f(x_1, y_1^{(1)})\right]$

$$= y_0 + \frac{h}{2}\left[\frac{1}{x_0 + y_0} + \frac{1}{x_0 + y_1^1}\right] = 1 + \frac{0.5}{2}\left[\frac{1}{0+1} + \frac{1}{0.5 + 1.375} = 1.3833\right]$$

Next, to compute the solution $y_2 = y(1) and\ let\ us\ consider\ x_0 = 0.5\,, y_0 = 1.3833$

$y_1^{(E)} = y_0 + hf(x_0, y_0) = y_0 + h\frac{1}{x_0+y_0} \quad = 1.3833 + \frac{0.5}{0.5+1.3833} = 1.6488$

First modification $y_1^{(1)} = y_0 + \frac{h}{2}\left[f(x_0, y_0) + f(x_1, y_1^{(E)})\right]$

$$= y_0 + \frac{h}{2}\left[\frac{1}{x_0 + y_0} + \frac{1}{x_0 + y_1^E}\right] = 1.3833 + \frac{0.5}{2}\left[\frac{1}{0.5 + 1.3833} + \frac{1}{1 + 1.6488}\right] =$$
$$= 1.6104$$

second modification $y_1^{(2)} = y_0 + \frac{h}{2}\left[f(x_0, y_0) + f(x_1, y_1^{(1)})\right]$

$$= y_0 + \frac{h}{2}\left[\frac{1}{x_0 + y_0} + \frac{1}{x_0 + y_1^1}\right] = 1.3833 + \frac{0.5}{2}\left[\frac{1}{0.5 + 1.3833} + \frac{1}{1 + 1.6104}\right]$$
$$= 1.6118$$

The required solution are y(0.5)=1.3833 and y(1)=1.6118

Example: 2

Using the modified Euler's method, solve the IVP $\frac{dy}{dx} = x + y^2, y(0) = 1$ at points x=0.1 in steps of length h=0.1. Carry out two modifications.

Solution:

$\frac{dy}{dx} = f(x, y) = x + y^2 \;\; x_0 = 0\,, y_0 = 1 \; taking \; h = 0.1$

$y_1^{(E)} = y_0 + hf(x_0, y_0) = y_0 + h(x_0 + y_0^2) \;\; = 1 + 0.1 * (0 + 1) = 1.1$

First modification $y_1^{(1)} = y_0 + \frac{h}{2}\left[f(x_0, y_0) + f(x_1, y_1^{(E)})\right]$

$$= y_0 + \frac{h}{2}[x_0 + y_0^2 + x_0 + ((y_1^E)^2)] = 1 + \frac{0.1}{2}[0 + 1 + (0.1 + 1.1^2)] = 1.1155$$

second modification $y_1^{(2)} = y_0 + \frac{h}{2}\left[f(x_0, y_0) + f(x_1, y_1^{(1)})\right]$

$$= y_0 + \frac{h}{2}[x_0 + y_0^2 + x_0 + ((y_1^1)^2)] = 1 + \frac{0.1}{2}[0 + 1 + (0.1 + 1.1155^2)] = 1.1172$$

Hence the required solution is y(01)=1.1172

Example :3

Using the modified Euler's method, solve the IVP $\frac{dy}{dx} = 1 + + \frac{y}{x}, y(1) = 2$ at points x=0.1 in steps of length h=0.2. Carry out three modifications.

Solution:

$\frac{dy}{dx} = f(x, y) = 1 + \frac{y}{x}\ x_0 = 1\ , y_0 = 2\ taking\ h = 0.2$

$$y_1^{(E)} = y_0 + hf(x_0, y_0) = y_0 + h\left[1 + \frac{y_0}{x_0}\right] \quad = 2 + 0.2(1 + \tfrac{2}{1}) = 2.6$$

First modification $y_1^{(1)} = y_0 + \frac{h}{2}\left[f(x_0, y_0) + f(x_1, y_1^{(E)})\right]$

$$= y_0 + \frac{h}{2}\left[1 + \frac{y_0}{x_0} + 1 + \frac{y_1^{(E)}}{x_0}\right] = 2 + \frac{0.2}{2}\left[1 + \frac{2}{1} + \left(1 + \frac{2.6}{1.2}\right)\right] = 2.6167$$

second modification $y_1^{(2)} = y_0 + \frac{h}{2}\left[f(x_0, y_0) + f(x_1, y_1^{(1)})\right]$

$$= y_0 + \frac{h}{2}\left[1 + \frac{y_0}{x_0} + 1 + \frac{y_1^{(1)}}{x_0}\right] = 2 + \frac{0.2}{2}\left[1 + \frac{2}{1} + \left(1 + \frac{2.6167}{1.2}\right)\right] = 2.6181$$

Third modification $y_1^{(3)} = y_0 + \frac{h}{2}\left[f(x_0, y_0) + f(x_1, y_1^{(2)})\right]$

$$= y_0 + \frac{h}{2}\left[1 + \frac{y_0}{x_0} + 1 + \frac{y_1^{(2)}}{x_0}\right] = 2 + \frac{0.2}{2}\left[1 + \frac{2}{1} + \left(1 + \frac{2.6181}{1.2}\right)\right] = 2.61812$$

Hence , the value of y at x=1.2 is 2.6182

Example:4

Using Modified Euler's method, find y(0.1) given $\frac{dy}{dx} = x^2 + y$ & $y = 1$ when x = 0 by taking h = 0.05. Perform two iterations in each step.

Solution:

$f(x, y) = x^2 + y,\ x_0 = 0,\ y_0 = 1$ & $h = 0.05$, Hence $f(x_0, y_0) = 0 + 1 = 1$

$\therefore y_1^{(0)} = y_0 + hf(x_0, y_0) = 1 + 0.05 \times 1 = 1.05$. Now $f(x_1, y_1^{(0)}) = 0.05^2 + 1.05 = 1.0525$

$\therefore y_1^{(1)} = y_0 + \frac{h}{2}\left[f(x_0, y_0) + f(x_1, y_1^{(0)})\right] = 1 + \frac{0.05}{2}\left[1 + 1.0525\right] = 1.0513.$

Now $f(x_1, y_1^{(1)}) = 0.05^2 + 1.0513 = 1.0538$

$\therefore y_1^{(2)} = y_0 + \frac{h}{2}\left[f(x_0, y_0) + f(x_1, y_1^{(1)}) \right] = 1 + \frac{0.05}{2}[1 + 1.0538] = 1.0513.$ Hence

$y_1 = y(0.05) = 1.0513$

Now $f(x_1, y_1) = 0.05^2 + 1.0513 = 1.0538$

$\therefore y_2^{(0)} = y_1 + hf(x_1, y_1) = 1.0513 + 0.05 \times 1.0538 = 1.1040$. Now

$f(x_2, y_2^{(0)}) = 0.1^2 + 1.1040 = 1.114$

$\therefore y_2^{(1)} = y_1 + \frac{h}{2}\left[f(x_1, y_1) + f(x_2, y_2^{(0)}) \right] = 1.0513 + \frac{0.05}{2}[1.0538 + 1.114] = 1.1055.$

Now $f(x_2, y_2^{(1)}) = 0.1^2 + 1.1055 = 1.1155$

$\therefore y_2^{(2)} = y_1 + \frac{h}{2}\left[f(x_1, y_1) + f(x_2, y_2^{(1)}) \right] = 1.0513 + \frac{0.05}{2}[1.0538 + 1.1155] = 1.1055.$

Hence $y_2 = y(0.1) = 1.1055.$

1.5 RUNGE- KUTTA METHOD

Consider the IVP $\frac{dy}{dx} = f(x, y), y(0) = y_0$ -----------------------(1)

To determine the solution of this problem at $x_n = x_0 + nh$ by using this method, where h is step length

According to the Euler's method , the solution at $x_1\ is\ y_0 + hf(x_0, y_0)$.

This can be rewritten as $y_1 = y_0 + K$

where

$$k_1 = hf(x_0, y_0)$$

$$k_2 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}\right)$$

$$k_3 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_2}{2}\right)$$

$$k_4 = hf(x_0 + h, y_0 + k_3)$$

$y_1 = y_0 + K = y_0 + \frac{k_1 + 2k_2 + 2k_3 + k_4}{6}$ is an approximate solution for the equation 1 at x_1 Is known as Runge-Kutta method of order four.

Example:1

Using RK method of fourth order , to find y(0.2), given that $\frac{dy}{dx} = \frac{y-x}{y+x}$ $and\ y(0) = 1\ take\ h = 0.2$

Solution

$\frac{dy}{dx} = f(x,y) = \frac{y-x}{y+x}$ $, x_{\,0} = 0, y_0 = 1\ and\ h = 0.2$ then $x_1 = x_0 + h = 0.2$

$k_1 = hf(x_0, y_0) = h\frac{y_0 - x_0}{y_0 + x_0}$ $=0.2\left(\frac{1-0}{1+0}\right) = 0.2$

$$k_2 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}\right) = h\left(\frac{y_0 + \frac{k_1}{2} - (\,x_0 + \frac{h}{2})}{y_0 + \frac{k_1}{2} + (\,x_0 + \frac{h}{2})}\right)$$

$= 0.2\ \left(\frac{1.1-0.1}{1.1+0.1}\right) = 0.16667$

$$k_3 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_2}{2}\right) = h\left(\frac{y_0 + \frac{k_2}{2} - (\,x_0 + \frac{h}{2})}{y_0 + \frac{k_2}{2} + (\,x_0 + \frac{h}{2})}\right)$$

$= 0.2\ \left(\frac{1+0.083335-0.1}{1.1+0.11+0.083335+0.1}\right) = \frac{0.196667}{1.183335} = 0.166197$

$$k_4 = hf(x_0 + h, y_0 + k_3) = h\ \left(\frac{y_0 + k_3 - (\,x_0 + h)}{y_0 + k_3 + (\,x_0 + h)}\right)$$

$$= 0.2\ \left(\frac{1.166197 - 0.2}{1.166197 + 0.2}\right) = \frac{0.1932394}{1.366197} = 0.14144$$

$$y_1 = y_0 + K = y_0 + \frac{k_1 + 2k_2 + 2k_3 + k_4}{6}$$

$$= 1 + \frac{0.2 + 2(0.16667) + 2(0.166197) + 0.14144}{6}$$

y_1=1.16786

Example :2

Using RK method of fourth order , to find y(0.2), given that $\frac{dy}{dx} = x^2 - y \; and \; y(0) = 1 \; take \; h = 0.1$

Solution

$\frac{dy}{dx} = f(x,y) = x^2 - y \,, x_{\,0} = 0, y_0 = 1 \; and \; h = 0.1$ then $x_1 = \; x_0 + h = 0.1$

$k_1 = hf(x_0, y_0) = h({x_0}^2 - y_0) = 0.1 * (0 - 1) = -0.1$

$$k_2 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}\right) = h\left(\left(x_0 + \frac{h}{2}\right)^2 - \left(y_0 + \frac{k_1}{2}\right)\right)$$

$= 0.1 \; \left(0 + \frac{0.1}{2}\right)^2 - \left(1 + \frac{-0.1}{2}\right) = 0.1(\,(0.05)^2 - 0.95) = -0.09475$

$$k_3 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_2}{2}\right) = h\left(\left(x_0 + \frac{h}{2}\right)^2 - \left(y_0 + \frac{k_2}{2}\right)\right)$$

$= 0.1(\,(0.05)^2 - (1 - 0.047375)) = -0.0950125$

$$k_4 = hf(x_0 + h, y_0 + k_3) = h\left((x_0 + h)^2 - (y_0 + k_3)\right)$$

$= 0.1(\,(0.05)^2 - (1 - 0.0950125)) = -0.0895$

$$y_1 = y_0 + K = y_0 + \frac{k_1 + 2k_2 + 2k_3 + k_4}{6} = 1 + \frac{-0.1 - 0.1895 - 0.190025 - 0.0895}{6}$$
$$= 0.90516$$

Example 3: Use fourth Runge –Kutta method to find at $x = 0.1$ given that $\frac{dy}{dx} = 3e^x$ +2y, y(0) and h=0.1

>> By data, f(x,y)=3e^x +2y, x_0 =0, y_0 =0,h=0.1

We shall compute $k_{1,} k_2, k_3, k_4$

$$k_1 = hf(x_{0,}y_0) = (0.1)\,f(0,0) = (0.1)[3e^0 + 2\times 0] = 0.3$$

$$k_2 = hf\big(x_0 + h\frac{h}{2}, y_0 + \frac{k_1}{2}\quad\big) = (0.1)\,f(0.05,0.15)$$

$$=(0.1)\quad \left[3e^{0.05} + 2(0.15)\right] = 0.3454$$

$$k_3 = hf\big(x_0 + \frac{h}{2}, y_0 + \frac{k2}{2}\quad\big) = (0.1)\,f(0.05,0.1727)$$

$$=(0.1)[3\,e^{0.05} + 2(0.1727]=0.3499$$

$$k_4 = hf(x_0 + h, y_0 + k_3) = (0.1)\,f(0.1,0.3499)$$

$$=(0.1)\left[3e^{0.1} + 2(0.3499)\quad\right] = 0.4015$$

$$y(x_0 + h) = y_0 + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

$$i.e., y(0.1) = 0 + \frac{1}{6}[0.3 + 2(0.3454) + 2(0.3499) + 0.4015$$

$$\therefore y(0.1) = 0.3487$$

Example 4:

Use fourth order Runge-kutta method to compute y(1.1) given that

$$\frac{dy}{dx} = xy^{\frac{1}{3}}\ ,\ y(1) = 1$$

Solution

>>By data $f(x,y) = xy^{\frac{1}{3}}$ $x_0 = 1$, $y_0 = 1$ and we need to compute y(1.1) implies $x_0 + h = 1.1 \therefore h = 0.1$

We shall compute k_1, k_2, k_3, k_4

$$k_1 = hf(x_{0,}y_0) = (0.1)\,f(1,1) = (0.1)\left[(1)(1)^{\frac{1}{3}}\right] = 0.1$$

$$k_2 = hf\left(x_0 + \frac{h}{2}, y_0 + {}^{k_1}\!\diagup_{2}\right) = (0.1)\,f(1.05,1.05)$$

$= \quad (0.1)[(1.05)(1.05)^{\frac{1}{3]}} = 0.1067$

$$k_3 = hf(x_0 + \frac{h}{2}, y_0 + \frac{k_2}{2} \quad) = (0.1)\, f(1.05, 1.05335)$$

$$= (0.1)[(1.05)91.05335)^{\frac{1}{3}}] = 0.1068$$

$$k_4 = hf(x_0 + h, y_0 + k3) = (0.1)\, f(1.1, 1.1068)$$

$= (0.1)\ [\ (1.1)(1.1068)^{\frac{1}{3}}] = 0.1138$

$y(x_0 + h) = y_0 + \frac{1}{6}$ [0.1+2(0.1067)+2(0.1068)+0.1138]

y(1.1)=1.1068

Example 5:

Using Runge Kutta method of fourth order solve $\frac{dy}{dx} + y = 2x$ **at** $x = 1.1$ **Given that y=3 at x = 1 initially**.

Solution:

>> We have $\frac{dy}{dx} = 2x - y, x_0 = 1, y_0 = 3$

$$f(x, y) = 2x - y, x_0 + h = 1.1 \therefore h = 0.1$$

$\frac{dy}{dx} = f(x, y) = 2x - y\ , x_{\ 0} = 1, y_0 = 3\ and\ h = 0.1$ then $x_1 = x_0 + h = 0.1$

$k_1 = hf(x_0, y_0) = h(2x_0 - y_0) = 0.1 * (2 - 3) = -0.1$

$k_2 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}\right) = 0.1hf(1.05, 2.95) =$

$= 0.1\ (2 * 1.05 - 2.95) = -0.085$

$k_3 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_2}{2}\right) = hf\ (1.05, 2.91425)$

$= 0.1(\ 2 * 1.05 - 2.91425) = -0.08575$

$k_4 = hf(x_0 + h, y_0 + k_3) = hf\ (1.1, 2.91425)$

$= 0.1(\ 2 * 1.1 - 2.91425) = -0.071425$

$$y_1 = y_0 + K = y_0 + \frac{k_1 + 2k_2 + 2k_3 + k_4}{6} = 1 + \frac{-0.1 - 0.085 - 0.08575 - 0.071425}{6}$$

$$= 2.9145$$

1.6 Milne's Method:

The method in which the construction of y_n involves the use of not only y_{n-1}but also predecessors are called multi step methods. In multi step methods two formulas are used in conjunction with each other- one for predicting the value of y_n and the other for correcting the predicted value of y_n.

Consider the IVP $\frac{dy}{dx} = f(x,y) - - - - - - - - - - - - - - - -(1)$

Let $y_0=y(x_0)$,$y_1=y(x_1)$,$y_2=y(x_2)$ and $y_3=y(x_3)$ be these known solutions.

Suppose we wish to determine the solution of equation (1) at the point $x_4 = x_3+h$.

Let us denote the required solution by $y_4=y(x_4)$.

First we predict the value of $y_4=y(x_4)$ by using Milne's predictor formula:

$$y_4^{(p)} = y_0 + \frac{4h}{3}(2f_1 - f_2 + 2f_3) \qquad ----(2)$$

$which\ can\ be\ computed\ with\ the\ help\ of\ the\ speciied\ x_0, x_1, x_2\ , x_3 and y_0, y_1, y_2, y_3$

Next we correct the value of y_4 by using the Milne's corrector formula:

$$y_4^{(c)} = y_2 + \frac{h}{3}\left(f_2 + 4f_3 + f_4^{(p)}\right)--------------------(3)$$

$$where f_4^{(p)} = f(x_4, y_4^{(p)})$$

$If\ we\ wish\ to\ have\ more\ accurate\ approximation\ for\ y\ , we\ employ\ the\ process\ repeatedly$

Example :1

Use Milne's predictor –corrector method to find the value of y at x=0.8, given $\frac{dy}{dx} = x - y^2$

Where y(0)=0 , y(0.2)=0.02 , y(0.4)=0.0795 , y(0.6)=0.1762. Apply corrector formula twice.

Solution:

$$Here\ f(x,y) = x - y^2, h = 0.2$$

X	Y	$f(x,y) = x - y^2$
x_0=0	y_0=0	$f_{0=0-0^2=0}$
X_1=0.2	y_1=0.02	$f_{1=0.2-0.02^2=0.1996}$
X_2=0.4	y_2=0.0795	$f_{2=0.4-0.0795^2=0.3937}$
X_3=0.6	y_3=0.1762	$f_{3=0.6-0.1762^2=0.5689}$

Now , Milne's Predictor formula yields the predicted value of y_4 as

$$y_4^{(p)} = y_0 + \frac{4h}{3}(2f_1 - f_2 + 2f_3) = 0 + \frac{4 * 0.2}{3}(2(0.1996) - 0.3937 + 2(0.5689)) \\ = 0.30488$$

$$f_4^{(p)} = f(x_4, y_4^{(p)}) = \ x_4 - {y_4^{(p)}}^2 = 0.8 - (0.30488)^2 = 0.7070$$

Now , the Milne's corrector formula gives a corrected value of y_4 as

$$y_4^{(c)} = y_2 + \frac{h}{3}\left(f_2 + 4f_3 + f_4^{(p)}\right) = 0.0795 + \frac{0.2}{3}(0.3937 + 4(0.5689) + 0.7070) \\ = 030459$$

To apply corrector formula second time we must use corrector as predictor and substitute in $f_4^{(p)}$

$$f_4^{(p)} = f(x_4, y_4^{(p)}) = \ x_4 - {y_4^{(p)}}^2 = 0.8 - (0.30459)^2 = 0.70723$$

$$y_4^{(c)} = y_2 + \frac{h}{3}\left(f_2 + 4f_3 + f_4^{(p)}\right) = 0.0795 + \frac{0.2}{3}(0.3937 + 4(0.5689) + 0.70723) \\ = 03046$$

This twice corrected value of y_4 is the required value of y at x=0.8.

Example :2

Use Milne's predictor –corrector method Given the differential equation , given 5x$\frac{dy}{dx} + y^2 - 2 = 0$ and the set of values of (x,y)given in the following table, find the value of y at x=4.5 using the Milne's method

Where y(4)=1 , y(4.1)=1.0049 , y(4.2)=1.0097 , y(4.3)=1.0143, y(4.4)=1.0187 . Apply corrector formula twice.

Solution:

$$Here\ f(x,y) = -\left[\frac{y^2-2}{5x}\right] = \left[\frac{2-y^2}{5x}\right]$$

X	Y	$f(x,y) = \left[\frac{2-y^2}{5x}\right]$
x_0=4	y_0=1	f_{0}=$\left[\frac{2-1.0097^2}{5(4.2)}\right]$=0.04669
X_1=4.1	y_1=1.0049	f_{1}=$\left[\frac{2-1.0097^2}{5(4.2)}\right]$=0.04669
X_2=4.2	y_2=1.0097	$\left[\frac{2-1.0097^2}{5(4.2)}\right] = 0.04669$
X_3=4.3	y_3=1.0143	$\left[\frac{2-1.0097^2}{5(4.2)}\right] = 0.04517$
X_4=4.4	Y_4=1.0187	$\left[\frac{2-1.0097^2}{5(4.2)}\right] = 0.04374$

Now , Milne's Predictor formula yields the predicted value of y_4 as

$$y_4^{(p)} = y_0 + \frac{4h}{3}(2f_1 - f_2 + 2f_3)$$

$$= 1.0049 + \frac{4*0.1}{3}(2(0.04669) - 0.04517 + 2(0.04374)) = 1.02299$$

$$f\left(X4, y_4^{(p)}\right) = \left[\frac{2-{y_4^{(p)}}^2}{5x}\right] = \left[\frac{2-1.02299^2}{5(4.5)}\right] = 0.042378$$

Now , the Milne's corrector formula gives a corrected value of y_4 as

$$y_4^{(c)} = y_2 + \frac{h}{3}\left(f_2 + 4f_3 + f_4^{(p)}\right) = 1.0143 + \frac{0.1}{3}(0.04517 + 4(0.04374) + 0.042378)$$
$$= 1.02305$$

To apply corrector formula second time we must use corrector as predictor and substitute in $f_4^{(p)}$

$$f_4^{(p)} = f(x_4, y_4^{(p)}) = \left[\frac{2 - 1.02305^2}{5(4.5)}\right] = 0.042372$$

$$y_4^{(c)} = y_2 + \frac{h}{3}\left(f_2 + 4f_3 + f_4^{(p)}\right) = 1.0143 + \frac{0.1}{3}(0.04517 + 4(0.04374) + 0.042372)$$
$$= 1.02305$$

This twice corrected value of y4 is the required value of y at x=4.5.

1.7 Adams- Bashforth Method:

The predictor and corrector formulas for this method is follows

$y_4^{(p)} = y_3 + \frac{h}{24}(55f_3 - 59f_2 + 37f_1 - 9f_0)$ ------------(1)

$$y_4^{(c)} = y_3 + \frac{h}{24}\left(f_1 - 5f_2 + 19f_3 + 9f_4^{(p)}\right) - - - - - - (2)$$

$$where f_4^{(p)} = f(x_4, y_4^{(p)})$$

Example :1

Using the Adams- Bashforth Method, solve the differential equation $\frac{dy}{dx} = x - y^2$ at x=0.8 , given that y(0)=0, y(0.2)=0.02 , y(0.4)=0.0795 , y(0.6)=0.1762.

Solution:

$Here\ f(x,y) = x - y^2, h = 0.2$

X	Y	$f(x,y) = x - y^2$
x_0=0	y_0=0	$f_{0=0-1=-1}$
X_1=0.2	y_1=0.02	$f_{1=}0.2 - 0.02^2 = 0.1996$
X_2=0.4	y_2=0.0795	$f_{2=}0.4 - 0.0795^2$ $= 0.39368$
X_3=0.6	y_3=0.1762	$f_{3=}0.6 - 0.1762^2$ $= 0.56895$

$$y_4^{(p)} = y_3 + \frac{h}{24}(55f_3 - 59f_2 + 37f_1 - 9f_0)$$

$= 0.1762 + \frac{0.2}{24}(55 * 0.56895 - 59 * 0.39368 + 37 * 0.1996 - 9 * 0) = 0.30495$ ------------(1)

$$where f_4^{(p)} = f(x_4, y_4^{(p)}) = \; x_4 - {y_4^{(p)}}^2 = 0.8 - 0.30495^2 = 0.70701$$

$$y_4^{(c)} = y_3 + \frac{h}{24}\left(f_1 - 5f_2 + 19f_3 + 9f_4^{(p)}\right) - - - - - - -(2)$$

$=0.1762 + \frac{0.2}{24}(0.1996 - 5 * 0.39368 + 19 * 0.56895 + 9 * 070701) = 0.30457$

To get a correction of this solution $y_4^{(p)} = 0.30457$

$$f_4^{(p)} = f(x_4, y_4^{(p)}) = \; x_4 - {y_4^{(p)}}^2 = 0.8 - 0.30457^2 = 0.70724$$

$y_4^{(c)} = 0.1762 + \frac{0.2}{24}(0.1996 - 5 * 0.39368 + 19 * 0.56895 + 9 * 070724) = 0.30459$

The required solution as y(0.8)=0.3046 correct to 4 decimal places.

Example :2

Using the Adams- Bashforth Method, solve the differential equation $\frac{dy}{dx} = x^2(1 + y)$ x at 1.4 . Carry out two corrections for the solution.

, given that y(1)=1, y(1.1)=1.233 , y(1.2)=1.548 , y(1.3)=1.979.

Solution:

$Here\ f(x, y) = x^2(1 + y)\ , h = 0.1$

X	Y	$f(x, y) = x^2(1 + y)$
x_0=1	y_0=1	$f_{0=1^2(1+1)=2}$
X_1=1.1	y_1=1.233	$f_1 = 2.702$
X_2=1.2	y_2=1.548	$f_2 = 3.669$
X_3=1.3	y_3=1.979	$f_3 = 5.035$

$$y_4^{(p)} = y_3 + \frac{h}{24}(55f_3 - 59f_2 + 37f_1 - 9f_0)$$

$= 1.979 + \frac{0.1}{24}(55 * 5.035 - 59 * 3.669 + 37 * 2.702 - 9 * 2) = 2.572$ ------------(1)

$$f_4^{(p)} = f(x_4, y_4^{(p)}) = \; x_4^2(1 + y_4^{(p)}) = 1.4^2(1 + 2.572) = 7.001$$

$$y_4^{(c)} = y_3 + \frac{h}{24}\left(f_1 - 5f_2 + 19f_3 + 9f_4^{(p)}\right) ------(2)$$

$= 1.979 + \frac{0.1}{24}(2.702 - 5 * 3.669 + 19 * 5.035 + 9 * 7.001) = 2.575$

To get a correction of this solution $y_4^{(p)} = 2.575$

$$f_4^{(p)} = f(x_4, y_4^{(p)}) = x_4^2(1 + y_4^{(p)}) = 1.4^2(1 + 2.575) = 7.007$$

$y_4^{(c)} = 0.1762 + \frac{0.1}{24}\left((2.702 - 5 * 3.669 + 19 * 5.035 + 9 * 7.007) = 2.5752\right)$ is the required solution as y(1.4)=2.5752 correct to 4 decimal places.

Example :3

Using the Adams- Bashforth Method, solve the differential equation $\frac{dy}{dx} + \frac{y}{x} = \frac{1}{x^2}$ at x=1.4 . Carry out two corrections for the solution.

, given that y(1)=1, y(1.1)=0.996, y(1.2)=0.986 , y(1.3)=0.972

Solution:

$Here\ f(x, y) = x^2(1 + y), h = 0.1$

X	Y	$f(x,y) = -\frac{y}{x} + \frac{1}{x^2}$
x_0=1	y_0=1	$f_{0=0}$
X_1=1.1	y_1=0.996	$f_1 = -0.079$
X_2=1.2	y_2=0.986	$f_2 = -0.12722$
X_3=1.3	y_3=0.972	$f_3 = -0.15598$

$$y_4^{(p)} = y_3 + \frac{h}{24}(55f_3 - 59f_2 + 37f_1 - 9f_0) ----- (1)$$

$= 0.972 + \frac{0.1}{24}(55 * -0.15598 - 59 * -0.12722 + 37 * -0.079 - 9 * 0) = 0.95535$

$$f_4^{(p)} = f(x_4, y_4^{(p)}) = -\frac{y_4^{(p)}}{x_4} + \frac{1}{x_4^2} = -\frac{0.95535}{1.4} + \frac{1}{1.4^2} = -0.1722$$

$$y_4^{(c)} = y_3 + \frac{h}{24}\left(f_1 - 5f_2 + 19f_3 + 9f_4^{(p)}\right) ------(2)$$

$= 0.972 + \frac{0.1}{24}(-0.079 - 5 * -0.12722 + 19 * -0.15598 + 9 * -0.1722) = 0.95552$

To get a correction of this solution let $y_4^{(p)} = 0.95552$

$$f_4^{(p)} = f(x_4, y_4^{(p)}) = -\frac{y_4^{(p)}}{x_4} + \frac{1}{x_4^2} = -\frac{0.95552}{1.4} + \frac{1}{1.4^2} = -0.17231$$

$y_4^{(c)} = 0.972 + \frac{0.1}{24}(-0.079 - 5 * -0.12722 + 19 * -0.15598 + 9 * -0.17231) = 0.95551$

is the required solution as y(1.4)= 0.95551 correct to 4 decimal places.

1.8 Course Outcomes: On completion of this course, students are able to:
1. Use appropriate single step and multi-step numerical methods to solve first order ordinary differential equations arising in flow data design problems

1.9 Exercises

1) Evaluate $y(0.1)$ correct to 6 places of decimals by Taylor's series method if $y(x)$ satisfies $y' = xy + 1,\ y(0) = 1$. [Ans: $y(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{8} + \frac{x^5}{15} + \frac{x^6}{48} + \cdots\cdots$ & y(0.1) = 1.1053]

2) Solve $y' = y^2 + x,\ y(0) = 1$ using Taylor's series method & compute $y(0.1)$ & $y(0.2)$ [Ans: $y(x) = 1 + x + \frac{3}{2}x^2 + \frac{4}{3}x^3 + \frac{17}{12}x^4 + \frac{31}{20}x^5 + \cdots\cdots$ and y(0.1) = 1.1165, y(0.2) = 1.2734]

3) Use Taylor's series method to find y at the point $x = 0.1$ & $x = 0.2$, given that $\frac{dy}{dx} = x^2 + y^2,\ y(0) = 1$

4) Given $\frac{dy}{dx} = \frac{y - x}{y + x}$ with boundary condition y = 1 when x = 0, find approximately y for x = 0.1 by Modified Euler's method [Ans: 1.0928]

5) Given that $\frac{dy}{dx} = x + y^2$ & $y = 1$ at $x = 0$. Find a approximate value of y at x = 0.5 by Modified Euler's method. [Ans: 2.2352]

6) Solve the differential equation $\frac{dy}{dx} = -xy^2,\ y = 2$ at $x = 0$, by Modified Euler's method & obtain y at x = 0.2 in two stages of 0.1 each. [Ans: 1.9227]

7) Using Runge – Kutta method of order 4, compute $y(0.2)$ & $y(0.4)$ from $10\frac{dy}{dx} = x^2 + y^2$, $y(0) = 1,$ taking h = 0.1 [Ans: 1.0207, 1.0438]

8) Use Runge – Kutta fourth order method to find y when x = 1.2 in steps of 0.1 given that $\frac{dy}{dx} = x^2 + y^2$ & $y(1) = 1.5$ [Ans: 2.5005]

9) Using Runge-Kutta method of order 4, compute y(0.2) for the equation, $y' = y - \frac{2x}{y}$, y(0) = 1.0 (Take h = 0.2) [Ans: 1.18323]

10) Using Runge – Kutta method of order 4, find y(0.2) for the equation $\frac{dy}{dx} = \frac{y-x}{y+x}$, $y(0) = 1.$ Take h = 0.2. [Ans: 1.1678]

11) If $\frac{dy}{dx} = 2e^x - y$, y(0) = 2, y(0.1) = 2.010, y(0.2) = 2.04 and y(0.3) = 2.09 find y(0.4) correct to four decimal places. By using Milne’s predictor-corrector method. (Use corrector formula twice). [Ans: 2.1621 & 2.2546]

12) Given $2\frac{dy}{dx} = (1 + x^2)y^2$ & $y(0) = 1,\ y(0.1) = 1.06,\ y(0.2) = 1.12,\ y(0.3) = 1.21.$ Evaluate $y(0.4)$ by Milne’s predictor – corrector method.[Ans: 1.2797]

13) Given $y^1 = x^2 - y,\ y(0) = 1$ & the starting values $y(0.1) = 0.90516,\ y(0.2) = 0.82127,$ $y(0.3) = 0.74918.$ Evaluate $y(0.4)$ using Adams – Bash forth method. [Ans: 0.6897]

14) Using Adams – Bash forth method, evaluate $y(1.4)$, if y satisfies $\frac{dy}{dx} + \frac{y}{x} = \frac{1}{x^2}$ & $y(1) = 1,$ $y(1.1) = 0.996,\ y(1.2) = 0.986,\ y(1.3) = 0.972.$ [Ans: 0.949]

1.10 Further Reading

1. *http://nptel.ac.in/courses.php?disciplineID=111*
2. *http://wwww.khanacademy.org/*
3. *http://www.class-central.com/subject/math*