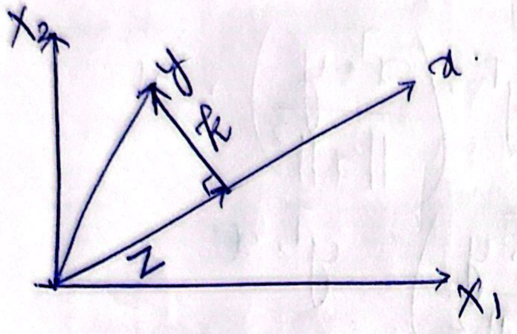
Data Mining

HW 1

a) $y^T \cdot z = \|z\|^2$.



→ Projected vector y on vector x ~~perpen~~ ~~dicularly~~.

→ The projection is z.

$y = k + z$.

$k = y - z$

$\rightarrow z = c \cdot x$ ($\because$ c is a scalar)

Replacing z value for equation 'k'.

$k = y - c \cdot x$.

→ Since z and k are seem to be perpen-dicular, their dot product, will be

$z^T \cdot k = 0$

$(c \cdot x)^T \cdot (y - cx) = 0$.

$c \cdot x^T y - c^2 x^T x = 0$ $(\because c^T = c)$

$\cancel{c}\, x^T y = c^{\cancel{2}} x^T x$

$\Rightarrow c = \dfrac{x^T \cdot y}{x^T \cdot x} = \dfrac{x^T \cdot y}{\|x\|^2}$

As we have value of c, we can find z

as. $z = c \cdot x$

$$= \left(\frac{x^T \cdot y}{\|x\|^2}\right) \cdot x.$$

$$\Rightarrow \quad y^T \cdot z = y^T \cdot \left(\frac{x^T \cdot y}{\|x\|^2}\right) \cdot x$$

$$= \left(\frac{x^T \cdot y}{\|x\|^2}\right) y^T \cdot x$$

$$= \left(\frac{y \cdot x}{\|x\|^2}\right) y \cdot x = \frac{(y \cdot x)^2}{\|x\|^2}$$

$$\rightarrow \text{R.H.S}, \quad \|z\|^2 = \left[\left(\frac{x^T \cdot y}{\|x\|^2}\right) \cdot x\right]^2$$

$$= \left(\frac{y \cdot x}{\|x\|^2}\right)^2 \cdot \|x\|^2.$$

$$= \frac{(y \cdot x)^2}{\|x\|^4} \cdot \|x\|^2.$$

$$= \frac{(y \cdot x)^2}{\|x\|^2} = y^T \cdot z$$

Scanned with CamScanner

$\therefore \quad y^T \cdot z = \|z\|^2$

$L.H.S = R.H.S.$

hence proved.

(b) Given equation as $(x-2)^2 + (y-5)^2 = 3^2$.

Center $= (2, 5)$ with a radius of 3.

Distance from Origin is $D^2 = x^2 + y^2$.

Consider F as Lagranging Function

$$F(x, y, \lambda) = x^2 + y^2 + \lambda\left((x-2)^2 + (y-5)^2 - 9\right)$$

$$\frac{\partial F}{\partial x} = 2 \cdot x + 0 + 2 \cdot \lambda (x-2) + 0.$$

$$= 2x + 2\lambda x - 4.$$

$$\frac{\partial F}{\partial x} = 0 \implies 2x + 2\lambda(x-2) = 0$$

$$x = \frac{2\lambda}{1+\lambda}.$$

$$\frac{\partial F}{\partial y} = 0.$$

$$\Rightarrow 2 \cdot y + 2\lambda \cdot (y-5) = 0.$$

$$\Rightarrow \quad y = \frac{5\lambda}{1+\lambda}$$

$$\frac{\partial F}{\partial \lambda} = 0.$$

$$\frac{\partial F}{\partial \lambda}\left(\lambda x^2 + \lambda 4 - 4x\lambda + \lambda y^2 + 25\lambda - 10y\lambda - 9\lambda\right) = 0.$$

$$x^2+y^2-4x-10y+20 = 0.$$

$$x^2+y^2-4x-10y = -20.$$

$$\left[\frac{2\lambda}{1+\lambda}\right]^2+\left[\frac{5\lambda}{1+\lambda}\right]^2-4\cdot\left(\frac{2\cdot\lambda}{1+\lambda}\right)-10\left(\frac{5\lambda}{1+\lambda}\right) = -20$$

$$p = \frac{\lambda}{1+\lambda}.$$

$$(2p)^2+(5p)^2-8p-50p = -20.$$

$$29p^2-58p+20 = 0.$$

$$p = \frac{-b\pm\sqrt{b^2-4ac}}{2a} = \frac{58\pm\sqrt{58^2-4(29)(20)}}{2(29)}$$

$$P_1 = \frac{58+\sqrt{(58)^2-4(29)(20)}}{58} \qquad P_2 = \frac{58-\sqrt{58^2-4(29)(20)}}{58}$$

$$= \frac{58+32.31}{58} \qquad = \frac{58-32.31}{58}$$

$$= 1.557 \qquad = 0.44$$

$$\frac{\lambda}{1+\lambda} = 1.557$$

$$\lambda = 1.557+1.557\lambda.$$

$$0.55\lambda = -1.55$$

$$\boxed{\lambda = \frac{-1.55}{0.55}}$$

$$\lambda = -2.81$$

$$x = \frac{2(-2.81)}{1-2.81}$$

$$= \frac{-5.62}{-1.81}$$

$$x = 3.10.$$

$$y = \frac{5(-2.81)}{-1.81}$$

$$y = 7.76$$

$\therefore (x, y) = (3.1, 7.7)$

This is the farthest point from the origin and also satisfies the circle equation.

(c) From projection formula, P_i denotes point

$$P_i = \left(\frac{P_i \cdot V_1}{V_1 \cdot V_1}\right) V_1 + \left(\frac{P_i \cdot V_2}{V_2 \cdot V_2}\right) V_2$$

$$V_1 \cdot V_1 = [2, 1, 2] \cdot [2, 1, 2] = 4 + 1 + 4 = 9$$

$$V_2 \cdot V_2 = [-1, 0, 1] \cdot [-1, 0, 1] = 1 + 0 + 1 = 2.$$

$$P_1 = [1, -1, 8]$$

$$\text{Projection of } P_1 = \frac{P_1 V_1}{9} [2, 1, 2] + \frac{P_1 V_2}{2} . \overset{-1,0,1}{\cancel{[2,1,2]}}$$

$$P_1 V_1 = [1, -1, 8]\ [2, 1, 2] = 17.$$

$$P_1 V_2 = [1, -1, 8]\ [-1, 0, 1] = 7.$$

$$\text{Proj on } P_1 = \frac{17}{9} [2, 1, 2] + \frac{7}{2} [-1, 0, 1]$$

$$= \left[\frac{34}{9}, \frac{17}{9}, \frac{34}{9}\right] + \left[\frac{-7}{2}, 0, \frac{7}{2}\right]$$

$$\text{Proj } P_1 = \left[\frac{5}{18}, \frac{17}{9}, \frac{131}{18}\right]$$

Scanned with CamScanner

Projec Squared distance $P_1 = ||P_1 - \text{Proj}(P_1)||^2$

$$= \left|\left|[1\ -1\ 8] - \left(\frac{5}{18}, \frac{17}{9}, \frac{131}{18}\right)\right|\right|^2$$

$$= 521/108 .$$

$$\text{Proje } P_2 = \left(\frac{P_2 V_1}{V_1 V_1}\right) V_1 + \left(\frac{P_2 \cdot V_2}{V_2 V_2}\right) V_2$$

$$P_2 V_1 = [4, 2, 1] \cdot [2, 1, 2] = 12$$

$$P_2 V_2 = [4, 2, 1] \cdot [-1, 0, 1] = -3$$

$$\text{Pro on } P_2 = \frac{12}{9}[2, 1, 2] + \frac{-3}{2}[-1, 0, 1]$$

$$= \left[\frac{8}{3}, \frac{4}{3}, \frac{8}{3}\right] + \left[\frac{3}{2}, 0, \frac{-3}{2}\right]$$

$$= \left[\frac{25}{6}, \frac{4}{3}, \frac{7}{6}\right]$$

$$\text{Pro on } P_3 = \left(\frac{P_3 V_1}{V_1 V_1}\right) V_1 + \left(\frac{P_3 V_2}{V_2 V_2}\right) V_2$$

$$P_3 V_1 = [0, 1, 5] \cdot [2, 1, 2] = 0 + 1 + 10 = 11$$

$$P_3 V_2 = [0, 1, 5] \cdot [-1, 0, 1] = 5 .$$

$$= \frac{11}{9}[2, 1, 2] + \frac{5}{2}[-1, 0, 1] = \left[\frac{22}{9}, \frac{11}{9}, \frac{22}{9}\right] + \left[\frac{-5}{2}, 0, \frac{5}{2}\right]$$

$$= \left[\frac{-1}{18}, \frac{11}{9}, \frac{89}{18}\right]$$

$$\text{Pro on } P_4 = \left(\frac{P_4 V_1}{V_1 V_1}\right) V_1 + \left(\frac{P_4 V_2}{V_2 V_2}\right) V_2$$

$$P_4 V_1 = [5, -2, -5] \cdot [2, 1, 2] = -2$$

$$P_4 V_2 = [5, -2, -5] \cdot [-1, 0, 1] = -10$$

Pro on $P_4 = \left(\frac{-2}{9}\right)(2,1,2] + -\frac{10}{2}[-1,0,1]$

$= \left[\frac{41}{9}, \frac{-2}{9}, \frac{-49}{9}\right]$

Pro on $P_5 = \left(\frac{P_5 V_1}{V_1 V_1}\right) V_1 + \left(\frac{P_5 V_2}{V_2 V_2}\right) V_2$

$P_5 V_1 = [-2, 0, -7] \cdot [2,1,2] = -18$

$P_5 V_2 = [-2, 0, -7] [-1,0,1] = 2-7 = -5$

Pro $P_5 = \frac{-18}{9}[2,1,2] + -\frac{5}{2}[-1,0,1] = \left[\frac{-3}{2}, -2, \frac{-13}{2}\right]$

Pro for $P_6 = \left(\frac{P_6 V_1}{V_1 V_1}\right) V_1 + \left(\frac{P_6 V_2}{V_2 V_2}\right) V_2$

$P_6 V_1 = [3\ 5, 3] \cdot [2,1,2] = 6+5+6 = 17$

$P_6 V_2 = [3, 5, 3] \cdot [-1, 0, 1] = -3+0+3 = 0$

$= \frac{17}{9}[2,1,2] + 0 = \left[\frac{34}{9}, \frac{17}{9}, \frac{34}{9}\right]$

Projection Matrix $= \begin{bmatrix} 5/18 & 17/9 & 131/18 \\ 25/6 & 4/3 & 7/6 \\ -1/18 & 11/9 & 89/18 \\ 41/9 & -2/9 & -49/9 \\ -3/2 & -2 & -13/2 \\ 34/9 & 17/9 & 34/9 \end{bmatrix}$

~~Mean Square error~~

$$\text{Actual} - \text{Projected} = \begin{bmatrix} 1 & -1 & 8 \\ 4 & 2 & 1 \\ 0 & 1 & 5 \\ 5 & -2 & -5 \\ -2 & 0 & -7 \\ 3 & 5 & 3 \end{bmatrix} - \begin{bmatrix} 5/18 & 17/9 & \frac{131}{18} \\ 25/6 & 4/3 & 7/6 \\ -1/18 & 11/9 & 89/18 \\ 41/9 & -2/9 & -49/9 \\ -3/2 & -2 & -13/2 \\ 34/9 & 17/9 & 34/9 \end{bmatrix} = \begin{bmatrix} 13/18 & -26/9 & 13/18 \\ -1/6 & 2/3 & -1/6 \\ 1/18 & -2/9 & 1/18 \\ 4/9 & -16/9 & 4/9 \\ -1/2 & 2 & -1/2 \\ -7/9 & -28/9 & -7/9 \end{bmatrix}$$

$$M.S.E = \frac{1}{6}\left[\left(\frac{13}{18} - \frac{26}{9} + \frac{13}{18}\right)^2 + \left(\frac{-1}{6} + \frac{2}{3} - \frac{1}{6}\right)^2 + \left(\frac{1}{18} + \frac{-2}{9} + \frac{1}{18}\right)^2 + \left(\frac{4}{9} - \frac{16}{9} + \frac{4}{9}\right)^2 + \left(\frac{-1}{2} + 2 - \frac{1}{2}\right)^2 + \left(\frac{-7}{9} - \frac{28}{9} - \frac{7}{9}\right)^2\right]$$

$$= \frac{1}{6}(25.77)$$

$$\boxed{M.S.E = 4.29}$$

(d) $A = L \cdot \Delta \cdot R^T$

$R^T = \Delta^{-1} L^T \cdot A$.

$$\Delta^{-1} = \frac{1}{\Delta}$$

$$= \begin{bmatrix} 1/11.18286 & 0 & 0 & 0 \\ 0 & 1/6.2933 & 0 & 0 \\ 0 & 0 & 1/3.6019 & 0 \\ 0 & 0 & 0 & 1/2.7147 \end{bmatrix}$$

$$= \begin{bmatrix} 0.0894 & 0 & 0 & 0 \\ 0 & 0.1588 & 0 & 0 \\ 0 & 0 & 0.2766 & 0 \\ 0 & 0 & 0 & 0.3683 \end{bmatrix}$$

for convention I am considering Δ^{-1} as 4x5 matrix.

$$\underset{4\times 5}{\Delta^{-1}} \cdot \underset{\substack{5\times 5 \\ 4\times 5}}{L^T} = \begin{bmatrix} 0.0894 & 0 & 0 & 0 \\ 0 & 0.1588 & 0 & 0 \\ 0 & 0 & 0.2766 & 0 \\ 0 & 0 & 0 & 0.3683 \end{bmatrix} \times$$

$$\begin{bmatrix} -0.4286 & 0.8338 & -0.2242 & -0.1063 & 0.2438 \\ -0.026 & -0.1394 & -0.4304 & -0.8429 & -0.2914 \\ -0.1034 & -0.332 & -0.7934 & 0.3315 & 0.3720 \\ 0.8939 & 0.3387 & -0.1709 & -0.0522 & 0.2332 \\ -0.0814 & -0.2441 & 0.3255 & -0.4069 & 0.8138 \end{bmatrix}$$

$\Delta^{-1} L^T =$

$$\begin{bmatrix} -0.0383 & 0.0745 & -0.0200 & -0.0095 & 0.0218 \\ -0.0004 & -0.0221 & -0.0683 & -0.1339 & -0.046 \\ -0.0287 & -0.0925 & -0.2202 & 0.0920 & 0.1035 \\ 0.329 & 0.1247 & -0.0629 & -0.0192 & 0.085 \end{bmatrix}$$

$R^T = (\Delta^{-1} L^T) A$

$$\begin{bmatrix} -0.0383 & 0.0745 & 0.0200 & -0.009 & 0.0218 \\ -0.0004 & -0.0221 & -0.0683 & -0.1339 & -0.046 \\ -0.0287 & -0.0925 & -0.2202 & 0.0920 & 0.1035 \\ 0.329 & 0.1247 & -0.0629 & -0.0192 & 0.085 \end{bmatrix}$$

$$\times \begin{bmatrix} 2 & 0 & -4 & 3 \\ -5 & 1 & 8 & 0 \\ 3 & -3 & 0 & 2 \\ 5 & 1 & 2 & 1 \\ 0 & 2 & 3 & 0 \end{bmatrix}$$

$$R^T = \begin{bmatrix} -0.557 & 0.1688 & 0.7962 & -0.1645 \\ -0.7648 & -0.0435 & -0.5823 & -0.2719 \\ 0.2044 & 0.8673 & -0.1306 & -0.4346 \\ -0.2502 & 0.4662 & -0.0997 & 0.842708 \end{bmatrix}$$