Date :- 12/01/2024

ROB-501 HW-9

(1) (a) $x_0 \in \text{span}\{y_1, \dots, y_p\}$

$\Rightarrow x_0 = \alpha_1 y_1 + \alpha_2 y_2 + \dots + \alpha_p y_p$

$\Rightarrow \langle x_0, y_i \rangle = c_i$

$\therefore \langle \alpha_1 y_1 + \alpha_2 y_2 + \dots + \alpha_p y_p, y_i \rangle = c_i$

$\langle \alpha_1 y_1, y_i \rangle + \langle \alpha_2 y_2, y_i \rangle + \dots + \langle \alpha_p y_p, y_i \rangle = c_i$

$\Rightarrow \alpha_1 \langle y_1, y_i \rangle + \alpha_2 \langle y_2, y_i \rangle + \dots + \alpha_p \langle y_p, y_i \rangle = c_i$

$$i=1,\dots,p \Rightarrow \underbrace{\begin{bmatrix} \langle y_1, y_1 \rangle & \langle y_2, y_1 \rangle & \dots & \langle y_p, y_1 \rangle \\ \langle y_1, y_2 \rangle & \langle y_2, y_2 \rangle & \dots & \langle y_p, y_2 \rangle \\ \vdots & \vdots & \ddots & \\ \langle y_1, y_p \rangle & \langle y_2, y_p \rangle & \dots & \langle y_p, y_p \rangle \end{bmatrix}} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix} = c_i$$

This is nothing but the gram-matrix & it is invertable if & only if $\det(G) \neq 0$ i.e The columns of the matrix is linearly independent.

$\Rightarrow \{y_1, \dots, y_p\}$ has to be linearly independent for ~~this to exist.~~ $\det(G) \neq 0$.

~~$\alpha_1 \neq \alpha_2 \neq \alpha_3 \neq \alpha_p$~~ $G\alpha = c_i \Rightarrow \alpha = G^{-1} c_i$

$\therefore$ all the co-efficients are unique ~~$x_0 = \xi$~~.

$\therefore x_0$ is obtained by the linear combination of $y_1 - \dots y_p \Rightarrow x_0$ is unique.

1(b)

1(b) Lemma 2:- let $M = (\text{Span}\{y_1, ---, y_p\})^\perp$. Then $V = x_0 + M$

in other words $x \in V$ if & only if $(x - x_0) \perp \text{Span}\{y_1, ---, y_p\}$

lets assume that $x \in V$ & w.k.T $x_0 \in V$

$\Rightarrow \langle x, y_i \rangle = c_i = \langle x_0, y_i \rangle$

Now lets consider the inner product $\langle x - x_0, y_i \rangle$

$\Rightarrow \langle x, y_i \rangle - \langle x_0, y_i \rangle$

w.k.T $\langle x, y_i \rangle = c_i$; & $\langle x_0, y_i \rangle = c_i$

$\therefore \langle x, y_i \rangle - \langle x_0, y_i \rangle = 0$

w.k.T $\langle x, y \rangle$ is 0 if $x \perp y$.

$\Rightarrow x - x_0 \perp y_i$

~~This is nothing but the Span $\{y^1, y^2, ---y^n\}$~~

~~$\Rightarrow x - x_0 \perp \text{Span}\{y^1, y^2, ---y^n\}$.~~

$\Rightarrow x - x_0 \in M$; Since M is the orthogonal complement of Span $\{y_1, ---y_p\}$

we can write $x = x_0 + m$ for some $m \in M$

Thus $x \in V$ if & only if x can be written as $x = x_0 + m$ for some $m \in M$

$\therefore$ we have shown that $V = x_0 + M$, this also confirms that any $x \in V$ satisfies $(x - x_0) \perp \text{Span}\{y_1, ---y_p\}$.

(i)

1(c) Expressing elements of V :- since w.k.t $V = x_0 + M$ any $v \in V$ can be written as $v = x_0 + m$ for some $m \in M$

where $M = (\text{span}\{y_1, \ldots y_p\})^{\perp}$

$$v^* \in V \Rightarrow v^* = \arg\min_{v \in V} \|v\|$$

Expressing v in terms of $m \in M$ we have:-

$$v = x_0 + m$$

$$\Rightarrow \cancel{v^* = \arg\min_{v \in V}} \|x_0 + m\|$$

$$\Rightarrow \cancel{m^* = \arg\min_{m \in M} \|x_0 + m\|}$$

↳ i.e we have changed the problem from v^* to m^*

now we have to find a way to exp v^* interms of m^* & hence we can use the projection theorem.

w.k.t given a vector space X & a closed subspace $M \subset X$ any element $x_0 \in X$ can be written as $x_0 = p + m^*$ where $p \in M^{\perp}$; $m^* \in M$

$\Rightarrow \|p\|$ is minimized when we choose m^* s.t m^* is the projection of x_0 onto M

So now the min. problem becomes.

$$\|x_0 - m^*\| = \inf_{m \in M} \|x_0 - m\| \Rightarrow \underline{v^* = x_0 - m^*}$$

$\therefore$ $v^* = x_0 - m^*$ is the unique element in V that has min-norm

$\Rightarrow$ v^* is orthogonal to M because m^* is chosen as the projection of x_0 onto M.

$\rightarrow$ This orthogonality does not imply that $v^* \perp V$ instead it only ensures that v^* is $\perp$ to the subspace M within X.

(2). let $\{y_1, \dots, y_p\}$ be a linearly independent set in X

& let $c_1, \dots, c_p$ be real constants.

$V = \{x \in X \mid \langle x, y_i \rangle = c_i \; 1 \le i \le p\}$. $v^* \in V$

S.T

$$\begin{bmatrix} \langle y_1, y_1 \rangle & \langle y_2, y_1 \rangle & \cdots & \langle y_p, y_1 \rangle \\ \langle y_1, y_2 \rangle & \langle y_2, y_2 \rangle & \cdots & \langle y_p, y_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle y_1, y_p \rangle & \langle y_2, y_p \rangle & \cdots & \langle y_p, y_p \rangle \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix}$$

lets consider lemma 1.

$x_0 \in \text{Span}\{y_1, \dots, y_p\}$ s.t $\langle x_0, y_i \rangle = c_i$

lets sub v^* instead of x_0.

$\langle v^*, y_i \rangle = c_i$ where $1 \le i \le p$.

v^* can be written as a lineal combination of $\{y_1, \dots, y_p\}$.

$\Rightarrow v^* = \sum_{i=1}^{p} \beta_i y_i$

$\langle \sum_{i=1}^{p} \beta_i y_i , y_i \rangle = c_i$ where $1 \le i \le p$

by expanding the inner product we get:-

$$\begin{bmatrix} \langle y_1, y_1 \rangle & \langle y_2, y_1 \rangle & \cdots & \langle y_p, y_1 \rangle \\ \langle y_1, y_2 \rangle & \langle y_2, y_2 \rangle & \cdots & \langle y_p, y_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle y_1, y_p \rangle & \langle y_2, y_p \rangle & & \langle y_p, y_p \rangle \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix}$$

(2) continued

lemma 1, lemma 2 states that

$\forall v \in V$:

$v = x^0 + m$

$m \in M$

$x^0 = \sum_{i=1}^{p} \beta_i y_i$

lemma 3 states that

$v^* = \underset{v \in V}{\operatorname{argmin}} \|v\|$ is unique, $v^* \perp M$

$\Downarrow$

$v^* = x^0 + m$; $\langle x^0 + m, m \rangle = \cancel{\langle x^0, m \rangle}^{0} + 0$. $(x^0 \perp m)$

$\Rightarrow \langle v^*, m \rangle = 0$

$\Rightarrow v^* = x^0 + m = \sum_{i=1}^{p} \beta_i y_i$

$\Rightarrow v^* = \sum_{i=1}^{p} \beta_i y_i$.

v^* is unique

$v^* = \underset{v \in V}{\operatorname{argmin}} \|v\|$

(3) (a) $\mu := \begin{bmatrix} \mathbb{E}\{x\} \\ \mathbb{E}\{y\} \end{bmatrix} = \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}$ & $\Sigma = \mathrm{cov}\left(\begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} x \\ y \end{bmatrix} \right)$

$$= \left[\begin{array}{c|c} P & PC^T \\ \hline CP & CPC^T + Q \end{array}\right]$$

$\Sigma_{x_1|x_2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$

$\mu_{1|2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$

$\Sigma_{11} = P \; ; \; \Sigma_{12} = PC^T$

$\Sigma_{21} = CP \; ; \; \Sigma_{22} = CPC^T + Q$

$$\mathbb{E}_{Y=y}(x) = \mu(1) + (PC^T)(CPC^T + Q)^{-1}(y - \mu(2))$$

$$\Rightarrow \cancel{E(x)_{Y=y} = \bar{x} + \frac{(PC^T)}{CPC^T+Q}(y - \bar{y})}$$

$$\boxed{E(x)_{Y=y} = \bar{x} + (PC^T)(CPC^T + Q)^{-1}(y - \bar{y})}$$

$$\boxed{\Sigma_{x|Y=y} = (P) - (PC^T)(CPC^T+Q)^{-1}(CP)}$$

(b) using Shus-complement we get.

$\Sigma_{CPC^T+Q} = P - PC^T(CPC^T + Q^{-1})CP = \Sigma_{x|Y=y}.$

$\Downarrow$

$\Sigma_{x|Y=y} = \Sigma_{CPC^T+Q}$

The way Σ_{CPC^T+Q} is solved using Shur Compliment is as follows.

$$\Sigma = \left[\begin{array}{c|c} P & PC^T \\ \hline CP & CPC^T + Q \end{array}\right] \Rightarrow \Sigma = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

$$\Sigma_{CPC^T+Q} = \Sigma_D = A - BD^{-1}C$$

$$\Rightarrow \Sigma_D = P - (PC^T)(CPC^T+Q)^{-1} \cdot (CP)$$

$$\Rightarrow \boxed{\Sigma_{CPC^T+Q} = P - (PC^T)(CPC^T+Q)^{-1} CP}$$

4(a) $\langle t,t\rangle \beta_i = 2$

$$\int_0^2 t^2 \cdot dt \quad \beta_i = 2$$

$$\frac{t^3}{3}\Big]_0^2 \quad \beta_i = 2$$

$$\Rightarrow \frac{8}{3}\beta_i = 2$$

$$\boxed{\beta_i = \frac{3}{4}} \quad \Rightarrow \begin{bmatrix}0\\ 3/4\\ 0\\ 0\end{bmatrix} = \beta_i$$

$$v^* = \frac{3}{4}t$$

4(b) $\cancel{\begin{bmatrix}\langle t,t\rangle & \langle t,\sin(\pi t)\rangle \\ \langle t,\sin(\pi t)\rangle & \end{bmatrix}\beta_i = \begin{bmatrix}2\\ \pi\end{bmatrix}}$

$$\begin{bmatrix}\langle t,t\rangle & \langle t,\sin(\pi t)\rangle \\ \langle t,\sin(\pi t)\rangle & \langle \sin(\pi t),\sin(\pi t)\rangle\end{bmatrix}\beta_i = \begin{bmatrix}2\\ \pi\end{bmatrix}$$

$$\beta_i = \begin{bmatrix} 3\pi^2/(2\pi^2-3) \\ \\ \pi(2\pi^2+3) \\ \hline (2\pi^2-3)\end{bmatrix} \approx \begin{bmatrix}1.7688\\ \\ 4.2677\end{bmatrix}$$

$$v^* = 1.7688t + 4.2677\sin(\pi t).$$

given $Ax = b$; $b \in \mathbb{R}^p$; $x \in \mathbb{R}^n$ where $n > p$.

5 a

$\|x\| = (x^T x)^{1/2} \Rightarrow$ inner product.

Thus the optimization problem is:

$$\hat{x} = \arg\min_{Ax=b} \|x\| = \arg\min_{Ax=b} (x^T x)$$

~~The gram-Matrix can be defined as $\tilde{A} = A$~~

The gram-matrix G associated with A is

$$G = \tilde{A} A^T = AA^T$$

The gram-matrix G is a PDF since the rows of A are linearly independent, G is invertible

$Ax = b$

$$A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix}$$

$$v = (a_i \cancel{A^T})^T = \cancel{A} a_i^T$$

$$Ax = b = \langle v, x \rangle \cancel{A} = b_i, \quad 1 \le i \le p$$

$$\|x\|^2 = x^T x \quad ; \quad x = A^T \lambda.$$

$$A(A^T \lambda) = b$$

$$G\lambda = b$$

$$\lambda = G^{-1} b$$

$$\lambda = (AA^T)^{-1} b.$$

$$\boxed{\hat{x} = A^T\lambda = A^T(AA^T)^{-1}b}$$

5(b) $\|x\| = (x^TQx)^{1/2}$

$$\hat{x} = \arg\min_{Ax=b} \|x\| = \arg\min_{Ax=b} (x^TQx)$$

$$\tilde{A} = AQ^{-1}$$

$$G = \tilde{A}\tilde{A}^T = AQ^{-1}A^T$$

$$A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix}$$

$$v_i = (a_iQ^{-1})^T = Q^{-1}a_i^T$$

$$\langle v, x\rangle_Q = v_i^TQx = a_i^Tx.$$

$\|x\|^2 = Q^{-1}A^T\lambda$; $Ax = b$ we get

$$A(Q^{-1}A^T\lambda) = b$$

$$G\lambda = b$$

$$\lambda = G^{-1}b = (AQ^{-1}A^T)^{-1}b$$

$$\boxed{\hat{x} = Q^{-1}A^T(AQ^{-1}A^T)^{-1}b}$$

(6) $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$

$v^1 = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = A^1 \therefore q^1 = \frac{v^1}{\|v^1\|} = \frac{1}{\sqrt{(1)^2+(3)^2+(5)^2}} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = \frac{1}{\sqrt{35}} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$

$v^2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} - \left\langle \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \frac{1}{\sqrt{35}} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \right\rangle \frac{1}{\sqrt{35}} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$

$v^2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} - \frac{1}{\sqrt{35}} \left\langle \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \right\rangle \frac{1}{\sqrt{35}} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$

$v^2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} - \frac{1}{\sqrt{35}} \left([2, 4, 6] \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \right) \frac{1}{\sqrt{35}} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$

$v^2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} - \frac{1}{35} (2 + 12 + 30) \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$

$v^2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} - \frac{44}{35} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$

$$r^2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} - \begin{bmatrix} \frac{44}{35} \\ \frac{44 \times 3}{35} \\ \frac{44}{7} \end{bmatrix} \Rightarrow = \begin{bmatrix} 2 - 44/35 \\ 4 - \frac{44 \times 3}{35} \\ 6 - \frac{44}{7} \end{bmatrix}$$

$$v^2 = \begin{bmatrix} \frac{26}{35} \\ \frac{8}{35} \\ -\frac{2}{7} \end{bmatrix}$$

$$q^2 = \frac{1}{\sqrt{\left(\frac{26}{35}\right)^2 + \left(\frac{8}{35}\right)^2 + \left(\frac{-2 \times 5}{7 \times 5}\right)^2}} \begin{bmatrix} \frac{26}{35} \\ \frac{8}{35} \\ -\frac{2}{7} \end{bmatrix}$$

$$q^2 = \frac{35}{\sqrt{676 + 64 + 100}} \begin{bmatrix} \frac{26}{35} \\ \frac{8}{35} \\ -\frac{2}{7} \end{bmatrix}$$

$$\Rightarrow q^2 = \frac{35}{\sqrt{840}} \begin{bmatrix} \frac{26}{35} \\ \frac{8}{35} \\ -\frac{10}{35} \end{bmatrix}$$

$$\Rightarrow q^2 = \frac{1}{\sqrt{840}} \begin{bmatrix} 26 \\ 8 \\ -10 \end{bmatrix}$$

$$r_{11} = \langle A^1, q^1 \rangle = \left\langle \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \frac{1}{\sqrt{35}} \right\rangle$$

$$= \frac{1}{\sqrt{35}} (35) = \sqrt{35}$$

$$r_{12} = \langle A^2, q^1 \rangle = \left\langle \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \frac{1}{\sqrt{35}} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \right\rangle$$

$$= \frac{1}{\sqrt{35}} \left[(2) + (12) + 30 \right] = \frac{44}{\sqrt{35}}$$

$$r_{21} = \langle A^1, q^2 \rangle = 0$$

$$r_{22} = \langle A^2, q^2 \rangle = \left\langle \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \frac{1}{\sqrt{840}} \begin{bmatrix} 26 \\ 8 \\ -10 \end{bmatrix} \right\rangle$$

$$r_{22} = \frac{1}{\sqrt{840}} \left[52 + 32 - 60 \right] = \frac{24}{\sqrt{840}}$$

QR decomposition of the matrix A is given as

$$A = \begin{bmatrix} \frac{1}{\sqrt{35}} & \frac{26}{\sqrt{840}} \\ \frac{3}{\sqrt{35}} & \frac{8}{\sqrt{840}} \\ \frac{5}{\sqrt{35}} & \frac{-10}{\sqrt{840}} \end{bmatrix} \begin{bmatrix} \sqrt{35} & \frac{44}{\sqrt{35}} \\ 0 & \frac{24}{\sqrt{840}} \end{bmatrix}$$

The (Q,R) decomposition in matlab gives the Result.

$$Q = \begin{bmatrix} -0.1690 & 0.8971 \\ -0.5071 & 0.2760 \\ -0.8452 & -0.3450 \end{bmatrix} \Rightarrow [Q,R] = q(A,0)$$

$$R = \begin{bmatrix} -5.9161 & -7.4374 \\ 0 & 0.8281 \end{bmatrix}$$

$[Q,R] = q(A)$
it gives a 3×3 matrix for Q.

$$Q = \begin{bmatrix} -0.1690 & 0.8971 & 0.4082 \\ -0.5071 & 0.2760 & -0.8165 \\ -0.8452 & -0.3450 & 0.4082 \end{bmatrix}$$

$$R = \begin{bmatrix} -5.9161 & -7.4374 \\ 0 & 0.8281 \\ 0 & 0 \end{bmatrix}$$

This is known as full Q.R decomposition where Q = 3×3

if we have to compute it by hand we have to find a unit vector q_3 which is orthogonal to q_1 & q_2 using Gram Schmidt. & the last row of R will be [0, 0] so that when we multiply Q * R we get the matrix A once again.