1. (a)

$\because x_0 \in \text{span}\{y_1 \cdots y_p\}$ $\therefore x_0 = \sum_{i=1}^{p} \alpha_i y_i$. $\alpha_i \in \mathbb{R}$

$\langle x_0, y_i \rangle = c_i \Rightarrow \langle \alpha_1 y_1 + \cdots \alpha_p y_p, y_i \rangle = c_i$

$\Rightarrow$ Normal equation : $G\alpha = \beta$, $[G]_{ij} = \langle y_i, y_j \rangle$, $\beta_j = \langle x_0, y_j \rangle$

$\because \{y_1 \cdots y_p\}$ are linear indep. , G is full rank and invertible

$\Rightarrow$ We can have unique solution $\alpha \Rightarrow x_0$ is unique. $\square$

(b)

Since $x_0 \in V \Rightarrow \langle x_0, y_i \rangle = c_i$

$\langle x - x_0, y_i \rangle = \langle x, y_i \rangle - \langle x_0, y_i \rangle = c_i - c_i = 0$, $1 \le i \le p$

$\Rightarrow x - x_0 \perp y_i \ (1 \le i \le p)$ thus $x - x_0 \perp \text{span}\{y_1 \ldots y_p\}$ $\square$

(c)

$v^* = \underset{v \in V}{\operatorname{argmin}} \|v\|$ and we know that $v \in V \Leftrightarrow v = x_0 - m$, $m \in M$

$\Rightarrow v^* = \underset{m \in M}{\operatorname{argmin}} \|x_0 - m\|$

$\Rightarrow x_0 - v^* \perp M \Rightarrow x_0 - v^* \perp (\text{span}\{y_1 \cdots y_p\})^{\perp}$

$\Rightarrow x_0 - v^* \in \text{span}\{y_1 \ldots y_p\}$

We know from lemma 2 . $v^* \in V \Leftrightarrow (x_0 - v^*) \perp \text{span}\{y_1 \ldots y_p\}$

$\therefore x_0 - v^* = 0 \Rightarrow x_0 = v^*$ $\square$

② From prob 1, we know $v^* = \underset{v \in V}{\arg\min} \|v\| = x_0$

$x_0 \in \operatorname{span}\{y_1 \cdots y_p\} = \sum_{i=1}^{p} \beta_i y_i$

plug into V, we get $\langle v^*, y_i \rangle = C_i$

$\Rightarrow \langle \beta_1 y_1 + \cdots \beta_p y_p, y_i \rangle = C_i$, we get normal eq. $G\beta = C$

$$G = \begin{bmatrix} \langle y_1, y_1 \rangle & \cdots & \langle y_p, y_1 \rangle \\ \vdots & & \vdots \\ \langle y_1, y_p \rangle & \cdots & \langle y_p, y_p \rangle \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} = \begin{bmatrix} C_1 \\ \vdots \\ C_p \end{bmatrix}$$

Since $\{y_1 \cdots y_p\}$ is linearly indep. G is full rank and invertible, we can get a set of unique β_i.

$\Rightarrow$ v^* is unique and minimum norm. $\square$

3.

(a)

$$\mu_{X|Y} = \mu_X + \Sigma_{XY}\Sigma_{YY}^{-1}(y-\mu_y)$$
$$= \bar{x} + PC^T(CPC^T+Q)^{-1}(y-\bar{y})$$

$$\Sigma_{X|Y} = \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}$$
$$= P - PC^T(CPC^T+Q)^{-1}CP$$

The result is the same as that in HW 8 prob 6.

(b) $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ $\quad M/D = A - BD^{-1}C$

$A = P,\ B = PC^T,\ C = CP,\ D = CPC^T+Q$

$\Rightarrow P - PC^T(CPC^T+Q)^{-1}CP$, which is the same as $\Sigma_{X|Y}$.

4.

(a)

Only 1 constraint, dimension of Gram matrix is 1×1

$f = \beta t$, $\langle f, t \rangle = 2$

$\Rightarrow [\langle t, t \rangle] \cdot \beta = C \Rightarrow 8/3\ \beta = 2 \Rightarrow \beta = 3/4$

$\Rightarrow f = 3/4\ t$

(b)

2 constraints $\Rightarrow \dim(G) = 2 \times 2$

$f = \beta_1 t + \beta_2 \sin(\pi t)$, $\langle f, t \rangle = 2$, $\langle f, \sin(\pi t) \rangle = \pi$

$$\begin{bmatrix} \langle t, t \rangle & \langle \sin(\pi t), t \rangle \\ \langle t, \sin(\pi t) \rangle & \langle \sin(\pi t), \sin(\pi t) \rangle \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 2 \\ \pi \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 1.7688 \\ 4.2677 \end{bmatrix}$$

$\Rightarrow f = 1.7688\, t + 4.2677 \sin(\pi t)$

5.

(a) Decompose $A = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$, let $v_i = a_i^T$

$$Ax = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} x = \begin{bmatrix} a_1 x \\ \vdots \\ a_n x \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

$$\langle v_i, x \rangle = v_i^T x = a_i x = b_i$$

Since we can write $Ax = b$ as $\langle v_i, x \rangle = b_i$, we can apply

Prob 2.

$$\exists\, x^* \text{ s.t. } x^* = \hat{x} = \underset{Ax=b}{\arg\min} \|x\|, \quad x^* = \sum_{i=1}^{n} \alpha_i v_i$$

α_i satisfy normal eq.

$$\begin{bmatrix} \langle v_1, v_1 \rangle & \cdots & \langle v_n, v_1 \rangle \\ \vdots & & \vdots \\ \langle v_1, v_n \rangle & \cdots & \langle v_n, v_n \rangle \end{bmatrix} \alpha = b \Rightarrow G\alpha = b$$

$$x^* = A^T \alpha$$

$$[G]_{ij} = \langle v_i, v_j \rangle = (a_i^T)^T a_j^T = a_i a_j^T$$

$$\Rightarrow G = AA^T$$

$$\therefore\ G\alpha = b,\ \alpha = G^{-1} b = (A^T A)^{-1} b \Rightarrow x^* = A^T \alpha$$

$$= A^T (AA^T)^{-1} b$$

∇

(b)

Decompose $A = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$, define $\tilde{A} = AQ^{-1}$ (since $Q > 0$, Q is invertible), and define $v_i := (a_i Q^{-1})^T = Q^{-1} a_i^T$

$$A x = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} x = \begin{bmatrix} a_1 x \\ \vdots \\ a_n x \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

Also, $\langle v_i, x \rangle = \langle Q^{-1} a_i^T, x \rangle = (Q^{-1} a_i^T)^T Q x = a_i Q^{-1} Q x = a_i x$
$= b_i$

$\therefore$ $Ax = b \iff \langle x, v_i \rangle = b_i$

Since we can write $Ax = b$ as $\langle x, v_i \rangle = b_i$, we can use the theorem in Prob. 2.

$\exists\, x^*$ s.t. $x^* = \hat{x} = \underset{Ax=b}{\arg\min} \|x\|$, $x^* = \sum_{i=1}^{n} \alpha_i v_i$ (α_i satisfy normal eq.)

$$\begin{bmatrix} \langle v_1, v_1 \rangle & \cdots & \langle v_n, v_1 \rangle \\ \vdots & & \vdots \\ \langle v_1, v_n \rangle & \cdots & \langle v_n, v_n \rangle \end{bmatrix} \alpha = b \implies G\alpha = b$$

$$x^* = \sum_{i=1}^{n} \alpha_i v_i = [v_1 \cdots v_n] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

$$= [Q^{-1} a_1^T \cdots Q^{-1} a_n^T] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

$$= Q^{-1} [a_1^T \cdots a_n^T] \alpha$$

$$= Q^{-1} A^T \alpha$$

$[G]_{ij} = \langle v_i, v_j \rangle = (Q^{-1} a_i^T)^T Q Q^{-1} a_j^T = a_i Q^{-1} Q Q^{-1} a_j^T$
$= a_i Q^{-1} a_j^T$

$\implies G = AQ^{-1}A^T$, $\alpha = G^{-1} b = (AQ^{-1}A^T)^{-1} b$

$\implies x^* = Q^{-1} A^T (AQ^{-1}A^T)^{-1} b$. $\square$

6

Hand calculation:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \quad V' = \frac{A_1}{\|A_1\|} = \begin{bmatrix} 0.1690 \\ 0.5071 \\ 0.8452 \end{bmatrix}$$

$$V^2 = \frac{A_2 - \langle A_2, V' \rangle \cdot V'}{\| A_2 - \langle A_2, V' \rangle \cdot V' \|} = \begin{bmatrix} 0.8971 \\ 0.2760 \\ -0.3450 \end{bmatrix}$$

$$Q = [V' V^2] = \begin{bmatrix} 0.1690 & 0.8971 \\ 0.5071 & 0.2760 \\ 0.8452 & -0.3450 \end{bmatrix}$$

$$R = \begin{bmatrix} \langle A_1, V' \rangle & \langle A_2, V' \rangle \\ 0 & \langle A_2, V^2 \rangle \end{bmatrix} = \begin{bmatrix} 5.9161 & 7.4374 \\ 0 & 0.8281 \end{bmatrix}$$

$[Q, R] = qr(A, 0)$

$$Q = \begin{bmatrix} -0.1690 & 0.8971 \\ -0.5071 & 0.2760 \\ -0.8452 & -0.3450 \end{bmatrix}, \quad R = \begin{bmatrix} -5.9161 & -7.4374 \\ 0 & 0.8281 \end{bmatrix}$$

$[Q, R] = qr(A)$

$$Q = \begin{bmatrix} -0.1690 & 0.8971 & 0.4082 \\ -0.5071 & 0.2760 & -0.8165 \\ -0.8452 & -0.3450 & 0.4082 \end{bmatrix}, \quad R = \begin{bmatrix} -5.9161 & -7.4374 \\ 0 & 0.8281 \\ 0 & 0 \end{bmatrix}$$

The Q, R calculated by hand is similar to that calculated by MATLAB function $qr(A, 0)$. The difference is that there are several extra negative signs in the Q, R from $qr(A, 0)$. The negative signs will be cancelled out when multiplying Q and R.

The Q, R calculated with qr(A) has an extra column in Q and an extra row in R. However, the extra row in R is all zero, so the extra column of Q will have no effect when multiplying Q and R.