CS 596 Midterm Xuenan Wang (XW336)

Problem 1.

(a) Since V denote the space of all polynomials of order up to n, We can assume a polynomial from V as:

$$p(x) = ax^n + bx^{n-1} + \cdots + cx + d, \text{ where } a, b, c, d \in \mathbb{R}.$$

To prove V is a vector space, we need to show 9 properties:

1) $p_1(x) \oplus (p_2(x) \oplus p_3(x)) = (p_1(x) \oplus p_2(x)) \oplus p_3(x)$

$p_1(x) \oplus (p_2(x) \oplus p_3(x)) = a_1x^n + b_1x^{n-1} + \cdots + c_1x + d_1 + (a_2x^n + b_2x^{n-1} + \cdots + c_2x + d_2 + a_3x^n + b_3x^{n-1} + \cdots + c_3x + d_3) = (a_1+a_2+a_3)x^n + (b_1+b_2+b_3)x^{n-1} + \cdots + (c_1+c_2+c_3)x + (d_1+d_2+d_3) = (p_1(x) \oplus p_2(x)) \oplus p_3(x)$

2) $p_1(x) \oplus p_2(x) = p_2(x) \oplus p_1(x)$

$p_1(x) \oplus p_2(x) = a_1x^n + b_1x^{n-1} + \cdots + c_1x + d_1 + a_2x^n + b_2x^{n-1} + \cdots + c_2x + d = p_2(x) \oplus p_1(x)$

3) $p(x) \oplus 0 = p(x)$

$p(x) \oplus 0 = ax^n + bx^{n-1} + \cdots + cx + d + 0x^n + 0x^{n-1} + \cdots + 0x + 0 = p(x)$

4) $-p(x) \oplus p(x) = 0$

$-p(x) \oplus p(x) = -ax^n - bx^{n-1} - \cdots - cx - d + ax^n + bx^{n-1} + \cdots + cx + d = 0$

5) $\alpha \odot p(x) \in V$

$\alpha \odot p(x) = \alpha \cdot ax^n + \alpha \cdot bx^{n-1} + \cdots + \alpha \cdot cx + \alpha \cdot d$, the outcome is also a polynomial of order up to n, which shows that $\alpha \odot p(x) \in V$.

6) $\alpha \odot (\beta \odot p(x)) = (\alpha \cdot \beta) \odot p(x)$

$\alpha \odot (\beta \odot p(x)) = \alpha \odot (\beta \cdot ax^n + \beta \cdot bx^{n-1} + \cdots + \beta cx + \beta d) = \alpha\beta ax^n + \alpha\beta bx^{n-1} + \cdots + \alpha\beta cx + \alpha\beta d = (\alpha \cdot \beta) \odot p(x)$

7) $1 \odot p(x) = p(x)$

$1 \odot p(x) = 1 \cdot ax^n + 1 \cdot bx^{n-1} + \cdots + 1 \cdot cx + 1 \cdot d = p(x)$

8) $\alpha \odot (p_1(x) \oplus p_2(x)) = (\alpha \odot p_1(x)) \oplus (\alpha \odot p_2(x))$

$\alpha \odot (p_1(x) \oplus p_2(x)) = \alpha \odot (a_1x^n + b_1x^{n-1} + \cdots + c_1x + d_1 + a_2x^n + b_2x^{n-1} + \cdots + c_2x + d_2) = \alpha a_1x^n + \alpha b_1x^{n-1} + \cdots + \alpha c_1x + \alpha d_1 + \alpha a_2x^n + \alpha b_2x^{n-1} + \cdots + \alpha c_2x + \alpha d_2 = (\alpha a_1x^n + \alpha b_1x^{n-1} + \cdots + \alpha c_1x + \alpha d_1) \oplus (\alpha a_2x^n + \alpha b_2x^{n-1} + \cdots + \alpha c_2x + \alpha d_2) = (\alpha \odot p_1(x)) \oplus (\alpha \odot p_2(x))$

9) $(\alpha + \beta) \odot p(x) = (\alpha \odot p(x)) \oplus (\beta \odot p(x))$

$(\alpha \odot p(x)) \oplus (\beta \odot p(x)) = (\alpha ax^n + \alpha bx^{n-1} + \cdots + \alpha cx + \alpha d) \oplus (\beta ax^n + \beta bx^{n-1} + \cdots + \beta cx + \beta d) = (\alpha+\beta)ax^n + (\alpha+\beta)bx^{n-1} + \cdots + (\alpha+\beta)cx + (\alpha+\beta)d = (\alpha+\beta) \odot p(x)$

Overall, all 9 properties satisfied. V is a vector space.

Addition properties are No.1 – No.4, multiplication properties are No.5 – No.9.

(b) V is finite dimensional, its dimension is $n+1$.

(c) $\{1, x, x^2, \cdots, x^n\}$

(d) Let S_1 denote the space of all polynomials $p_1(x)$ of order up to $n-1$; let S_2 denote the space of all polynomials $p_2(x)$ of order up to $n-2$; let S_3 denote the space of all polynomials $p_3(x)$ of order up to $n-3$. Then we have three subsets S_1, S_2, S_3 of vector space V.

(e) Let $p = [x^0 \ x^1 \ x^2 \ \cdots \ x^n]$, $C = [p_0 \ p_1 \ p_2 \cdots p_n]$

$$p(x) = \langle p, C \rangle = p \cdot C^T = [x^0 \ x^1 \ x^2 \cdots x^n] \begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix} = \sum_{i=0}^{n} x^i p_i$$

p is a base of V, and C can be writen as $[p_0x^0 \ p_1x^0 \ p_2x^0 \cdots p_nx^0]$ and it's also a polynomial of order up to n, which indicates that $C \in V$. And since this mapping relation is one-to-one, we have a inner product for V.

$$\|p(x)\|_l = \left(\sum_{i=0}^{l} p_i x^i\right)^{\frac{1}{l}}, \quad 1 \le l \le n.$$

Problem 2.

(a). Since Q is a real symmetric matrix, its eigenvectors x_i are orthogonal to each other.

$\because Qx_1 = \lambda_1 x_1, \quad Qx_2 = \lambda_2 x_2$

$\therefore x_2^T Q x_1 = \lambda_1 x_2^T x_1$

$\because Q = Q^T$

$\therefore x_1^T Q x_2 = \lambda_1 x_1^T x_2 . \quad x_1^T Q x_2 = \lambda_2 x_1^T x_2$

$\therefore (\lambda_1 - \lambda_2) x_1^T x_2 = 0$

$\because \lambda_1 \ne \lambda_2$

$\therefore x_1^T x_2 = 0$

We consider all eigenvectors as a set of base, denote as b_i, $i = 1 \cdots n$. and $(b_i, b_j) = 0$, if $i \ne j$. $Qb_i = \lambda_i b_i$

For vector x, we can say that: $x^* = \sum_{i=1}^{n} \alpha_i b_i^*$, where α_i is a scalar.

$\therefore (x^*)^T x = \sum_{i=1}^{n} \alpha_i^2 b_i^* b_i$

$$\begin{aligned}(x^*)^T Q x = \langle x^*, Qx \rangle &= \langle \sum_{i=1}^{n} \alpha_i b_i^*, Q \sum_{i=1}^{n} \alpha_i b_i \rangle \\ &= \langle \sum_{i=1}^{n} \alpha_i b_i^*, \sum_{i=1}^{n} \alpha_i Q b_i \rangle \\ &= \langle \sum_{i=1}^{n} \alpha_i b_i^*, \sum_{i=1}^{n} \alpha_i \lambda_i b_i \rangle \\ &= \sum_{i=1}^{n} \lambda_i \alpha_i^2 b_i^* b_i\end{aligned}$$

$$\frac{(x^*)^T Q x}{(x^*)^T x} = \frac{\sum_{i=1}^{n} \lambda_i a_i^2 b_i^* b_i}{\sum_{i=1}^{n} a_i^2 b_i^* b_i} = \lambda_i$$

$\because \rho_1 \geq \lambda_i \geq \rho_k$

$\therefore \rho_1 \geq \frac{(x^*)^T Q x}{(x^*)^T x} \geq \rho_k$

(b). $\sigma_1^2 \geq \frac{(x^*)^T A^T A x}{(x^*)^T x} \geq \sigma_k^2$

$$\frac{(x^*)^T A^T A x}{(x^*)^T x} = \frac{(Ax^*)^T Ax}{(x^*)^T x}$$

$\because Ax^* = \lambda^* x^*$,

we have $\frac{(x^*)^T A^T A x}{(x^*)^T x} = \frac{(\lambda^* x^*)^T \lambda x}{(x^*)^T x} = \frac{\lambda^* \lambda (x^*)^T x}{(x^*)^T x} = \lambda^* \lambda$

$\therefore \sigma_1^2 \geq \lambda^* \lambda \geq \sigma_k^2$

$\therefore \sigma_1 \geq |\lambda| \geq \sigma_k$

Problem 3.

(a) Assume λ is an eigenvalue of P and v is the corresponding eigenvector.

$$\lambda^2 v = P^2 v = Pv = \lambda v$$

since $v \neq 0$, $\lambda^2 = \lambda$

$\therefore \lambda_1 = 0$, $\lambda_2 = 1$

Therefore eigenvalue of P are either 0 or 1.

(b) $(I-P)^2 = I - P - P + P^2$

Since P is projection, we have $P = P^2$.

$(I-P)^2 = I - P$

So we have that $(I-P)^2 = I-P$, which is proof of $(I-P)$ is also a projection.

(c) $(x-Px)^T \cdot Px = (x^T - x^T P^T) Px = x^T Px - x^T P^T Px = x^T Px - x^T P^2 x$
$= x^T Px - x^T Px = 0.$

$\therefore$ $x-Px$ and Px are orthogonal.

(d) $P = A(B^T A)^{-1} B^T$

$$P^2 = A(B^T A)^{-1} B^T \cdot A(B^T A)^{-1} B^T = A(B^T A)^{-1} B^T = P$$

Therefore, we show that $P = A(B^T A)^{-1} B^T$ is a projection matrix.

To make this P well defined, we need $\det(B^T A) \neq 0$, m and n are positive integers.

When $P = A(B^T A)^{-1} B^T$ is symmetric matrix, we know that $P^T = P$ and P is an orthogonal projection.

(e) Let $\hat{b} = Ax$.

$\min \|\hat{b} - b\|^2 = \min \|Ax - b\|^2$

get the derivative of $(Ax-b)^T (Ax-b)$:

$$2A^T Ax - 2A^T b = 0.$$

$$A^T Ax = A^T b$$

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

$$\hat{b} = Ax = A(A^T A)^{-1} A^T b$$

$$P = A(A^T A)^{-1} A^T$$

$$P^2 = A(A^T A)^{-1} A^T A (A^T A)^{-1} A^T = A(A^T A)^{-1} A^T = P$$

Therefore, P is orthogonal projection.

Problem 4.

(a) Consider $x-\hat{x}$, where $\hat{x} = az + bw$.

Based on orthogonal principle, we know that:

$$\langle x-\hat{x}_*, w\rangle = 0 \quad \text{and} \quad \langle x-\hat{x}_*, z\rangle = 0$$

Since $\hat{x}_* = a_* z + b_* w$

$$\therefore \langle x - a_* z - b_* w, w\rangle = 0 \quad \text{and} \quad \langle x - a_* z - b_* w, z\rangle = 0$$

$\langle x - a_* z - b_* w, w\rangle = \langle x, w\rangle + \langle -a_* z - b_* w, w\rangle = \langle x, w\rangle - \langle a_* z + b_* w, w\rangle = 0$

$$\therefore \langle x, w\rangle = a_* \langle z, w\rangle + b_* \langle w, w\rangle$$

$\langle x - a_* z - b_* w, z\rangle = \langle x, z\rangle + \langle -a_* z - b_* w, z\rangle = \langle x, z\rangle - \langle a_* z + b_* w, z\rangle = 0$

$$\therefore \langle x, z\rangle = a_* \langle z, z\rangle + b_* \langle w, z\rangle$$

$$\begin{bmatrix} \langle x, w\rangle \\ \langle x, z\rangle \end{bmatrix} = \begin{bmatrix} \langle z, w\rangle & \langle w, w\rangle \\ \langle z, z\rangle & \langle w, z\rangle \end{bmatrix} \begin{bmatrix} a_* \\ b_* \end{bmatrix}$$

$$\therefore a_* = \frac{E[xw] - b_* E[w^2]}{E[zw]} \qquad b_* = \frac{E[xz]E[zw] - E[xw]E[z^2]}{E^2[wz] - E[w^2]E[z^2]}$$

$$\therefore a_* = \frac{E[wz]E[wx] - E[xz]E[w^2]}{E^2[wz] - E[z^2]E[w^2]}$$

(b) $\hat{x}_* = a_* z + b_* w = \dfrac{E[wz]E[wx] - E[xz]E[w^2]}{E^2[wz] - E[z^2]E[w^2]} \cdot z + \dfrac{E[xz]E[zw] - E[xw]E[z^2]}{E^2[wz] - E[w^2]E[z^2]} \cdot w$

Accroding to the question, we need $\min \|x-\hat{x}\|$

$$\|x-\hat{x}_*\|^2 = \langle x-\hat{x}_*, x-\hat{x}_*\rangle = \langle x, x-\hat{x}_*\rangle = \langle x, x\rangle - \langle x, \hat{x}_*\rangle$$

$= \langle x, x\rangle - \langle x, a_* z + b_* w\rangle = \langle x, x\rangle - [a_* \langle x, z\rangle + b_* \langle x, w\rangle]$

$= E[x^2] - (a_* E[xz] + b_* E[wx])$

$$\therefore \|x-\hat{x}_*\| = \sqrt{\|x-\hat{x}_*\|^2} = \sqrt{E[x^2] - (a_* E[xz] + b_* E[wx])}$$