# HW3 Theory

Saturday, February 13, 2021 11:21 AM

1) **1.** If $P$ is an orthogonal projector, then $I - 2P$ is unitary. Prove this algebraically and give a geometric interpretation.

<u>Proof</u>: Let $P$ be an orthogonal projector, i.e.

$P^2 = P = P^*$.

Then, $(I-2P)^*(I-2P) = (I^* - 2^*P^*)(I-2P)$

$= I^2 - 2PI - 2P^*I + 4P^*P$

$= I - 2P - 2P + 4P^2$

$= I - 2P^2 - 2P^2 + 4P^2$

$= I.$

Hence, $(I-2P)^{-1} = (I-2P)^*$, meaning

$I-2P$ is unitary. ▨

Geometrically, we know that $(I-P)v$ is a vector orthogonal to range$(P)$. Hence, $(I-2P)v$ is the reflection of $v$ through range$(P)$.

2) **2.** Given $A \in \mathbb{R}^{m\times n}$ with $m \geq n$. Show that $A^*A$ is nonsingular if and only if $A$ has full rank.

<u>Proof</u>: $(\Rightarrow)$ Let rank$(A^*A) = n$. Then, null$(A^*A) = \{0\}$,

i.e. $A^*Ax \neq 0 \quad \forall x \neq 0$. Therefore,

$A^*(Ax) \neq 0$, meaning $Ax \neq 0$. Thus rank$(A) = n$.

$(\Leftarrow)$ Assume rank$(A) = n$. Then, null$(A) = \{0\}$, i.e.

$Ax \neq 0 \quad \forall x \neq 0$. Left multiplying by $A^*$ gives us

$A^*(Ax) \neq A^* \cdot 0 = 0$. Thus $A^*Ax \neq 0 \quad \forall x \neq 0$,

meaning null$(A^*A) = \{0\} \Rightarrow$ rank$(A^*A) = n$

$\Rightarrow A^*A$ is nonsingular.

3) **3.** Consider the matrices

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{1}$$

Answer the following questions by hand calculation.

**(a)** What is the orthogonal projector $P$ onto range$(A)$, and what is the image under $P$ of $v = [1,2,3]^T$?

**(b)** Do the same for $B$.

<u>Solution</u>:

a) The columns of $A$ are linearly independent , Solution:

a) The columns of $A$ are linearly independent and orthogonal. Thus, normalizing the first column of $A$ yields an orthogonal matrix $Q$. Hence, from lecture, we can construct $P$ by computing

$$P = QQ^T.$$

Computation:

$$Q = \begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & 1 \\ 1/\sqrt{2} & 0 \end{bmatrix}$$

$$\Rightarrow P = QQ^T = \begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & 1 \\ 1/\sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 0 & 1\sqrt{2} \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix}.$$

$$P_v = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} + \frac{3}{2} \\ 2 \\ \frac{1}{2} + \frac{3}{2} \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}.$$

b) The columns of $B$ are linearly independent, but not orthogonal. Hence, we apply Graham-Schmidt.

Let $B = [b_1 \mid b_2]$.

Set $v_1 = b_1$. Then,

$$v_2 = b_2 - \frac{\langle b_2, v_1 \rangle}{\|v_1\|^2} v_1$$

$$v_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \frac{2}{(\sqrt{2})^2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}.$$

$$\langle v_1, v_2 \rangle = [1, 0, 1] \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = 1 + 0 - 1 = 0 \checkmark$$

Hence, an orthonormal basis for the range$(B)$ is

Hence, an orthonormal basis for the range(B) is

$$V = \left\{ \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ -1/\sqrt{3} \end{bmatrix} \right\}.$$ Thus,

let $Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{3} \\ 0 & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{3} \end{bmatrix}$. Then, $P = QQ^T$ gives us

$$P = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{3} \\ 0 & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}+\frac{1}{3} & \frac{1}{3} & \frac{1}{2}-\frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{2}-\frac{1}{3} & \frac{-1}{3} & \frac{1}{2}+\frac{1}{3} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{5}{6} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{3} & \frac{-1}{3} \\ \frac{1}{6} & \frac{-1}{3} & \frac{5}{6} \end{bmatrix}.$$ Therefore, $P_v = \begin{bmatrix} \frac{5}{6} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{3} & \frac{-1}{3} \\ \frac{1}{6} & \frac{-1}{3} & \frac{5}{6} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

$$= \begin{bmatrix} \frac{5}{6} + \frac{4}{6} + \frac{3}{6} \\ \frac{1}{3} + \frac{2}{3} - \frac{3}{3} \\ \frac{1}{6} - \frac{4}{6} + \frac{15}{6} \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}.$$

4)

**4.** Let $P \in \mathbb{R}^{m \times m}$ be a nonzero projector.

**(a)** Show that $||P||_2 \geq 1$, with equality if and only if $P$ is an orthogonal projector.

**(b)** If $P$ is an orthogonal projector, then $P$ is positive semi-definite with its eigenvalues are either zero or 1.

a) Since $P$ is nonzero, $\exists v \neq 0$, $\|v\|_2 = 1$, s.th

$$P_v = v.$$

Recall $\|P\|_2 = \sup_{\|x\|_2 = 1} \|Px\|_2.$

Thus, $\|Pv\|_2 = \|v\|_2 = 1 \leq \sup_{\|x\|_2 = 1} \|Px\| = \|P\|_2$ by

definition of supremum. Hence $\|P\|_2 \geq 1$.

Claim: $\|P\|_2 = 1$ if and only if $P$ is an orthogonal projector.

($\Leftarrow$): Assume $P$ is an orthogonal projector and consider its SVD:

$$P = U \Sigma V^T.$$ Then,

we have that

$$P^TP = P^2 = P = (U\Sigma V^T)^T(U\Sigma V^T)$$
$$= (V\Sigma U^T)(U\Sigma V^T)$$
$$= V\Sigma^2 V^T.$$

Thus, since $P^2 = P$, we can equate

$V\Sigma^2V^T = U\Sigma V^T$. Recall $\Sigma$ contains the singular values of $P$ (and thus $P^2$) along the diagonal, with $\sigma_1$ being the largest. Since orthogonal matrices are norm-preserving, we have

$$\|P\|_2 = \|P^2\|_2 = \|\Sigma^2\|_2 = \|\Sigma\|_2 \Rightarrow \sigma_1^2 = \sigma_1.$$

Thus $\sigma_1(\sigma_1 - 1) = 0$, implying $\sigma_1 = 0$ or $\sigma_1 = 1$.

Since $P$ is nonzero by assumption, $\sigma_1 = 1$.

Thus, $\|P\|_2 = 1$.

($\Rightarrow$) Now, assume $\|P\|_2 = 1$. We want to show $P$ is an orthogonal projector. Since $P$ is a nonzero projector, i.e. $P^2 = P$, it remains to show that $P$ is symmetric.

Unsure of how to proceed.

$$\sigma_i = \sqrt{\lambda_i(P^TP)}$$

$$\sigma_{max}(P) = \sigma_{max}(P^2) = \sqrt{\lambda_i(P^TP)} = \sqrt{\lambda_i(P^{2T}P^2)} = 1$$

b)

Let $P$ be a nonzero projector.

D1

Let $P$ be a nonzero projector.

Then, $P^2 = P$. Let $x$ be an eigenvector of $P$ with eigenvalue $\lambda$.

Then, $Px = \lambda x$, and $P^2 x = \lambda x$, since $P^2 = P$.

But also, left multiplying, we have

$$P^2 x = Px = \lambda Px = \lambda(\lambda x) = \lambda^2 x.$$

Hence, we have $Px = \lambda^2 x$ and also $Px = \lambda x$. Thus, $\lambda^2 x = \lambda x$, so

$$\lambda^2 x - \lambda x = \lambda(\lambda x - x) = 0.$$

Thus, $\lambda = 0$ or $\lambda x = x \Rightarrow \lambda = 1$.

Hence, if we assume $P$ is also an orthogonal projector, then $P^T = P$. Then, since $\lambda = 0$ or $\lambda = 1$ and $P$ is symmetric, by definition $P$ is positive semi-definite.

5) **5.** Consider the matrices in Problem 3.

**(a)** Using any method you prefer, determine a reduced QR factorization $A = \hat{Q}\hat{R}$ and a full QR factorization $A = QR$ by hand calculation.

**(b)** Do the same for $B$.

From 3),

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

We want to write $A = \hat{Q}\hat{R}$, $\hat{Q}$ has orthonormal columns, $\hat{R}$ upper triangular.

Computation:

$$\hat{Q} = \begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & 1 \\ 1/\sqrt{2} & 0 \end{bmatrix}.$$

We know $\hat{Q}^T A = \hat{R}$,

$$\Rightarrow \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 1 \end{bmatrix} = \hat{R}.$$

Thus $A = \hat{Q}\hat{R} = \begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & 1 \\ 1/\sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 1 \end{bmatrix}$ (Reduced Rank QR fac.)

Thus $A = QR = \begin{array}{c} \begin{bmatrix} 0 & 1 \\ 1/\sqrt{2} & 0 \end{bmatrix} \\ \quad q_1 \quad\; q_2 \end{array} \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 1 \end{bmatrix}$ (Reduced Rank QR fac.)

To produce the full-rank we need to add another orthonormal column, $q_3 = \frac{v_3}{\|v_3\|}$ to $\hat{Q}$. Thus, the following must hold

$\langle q_1, v_3 \rangle = 0$, ① $\Rightarrow \frac{1}{\sqrt{2}} v_{13} + \frac{1}{\sqrt{2}} v_{33} = 0$

$\langle q_2, v_3 \rangle = 0$, ② $\Rightarrow v_{23} = 0$

$\|q_3\| = 1$.

① $\Rightarrow$ $v_{13} = -v_{33}$. Let $v_{13} = 1$, so $v_{33} = -1$.

Then, $v_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$. Thus, $q_3 = \frac{v_3}{\|v_3\|} = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}$.

Hence our full QR factorization is

$$A = \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

b) We have

$B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$. From problem 3), an orthonormal basis for the column space of B is

$V = \left\{ \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ -1/\sqrt{3} \end{bmatrix} \right\}$. $\leftarrow$ columns of $\hat{Q}$.

For the reduced QR factorization of B, we have:

$$B = \hat{Q}\hat{R} \Rightarrow \hat{Q}^T B = \hat{R}$$

$$\Rightarrow \hat{R} = \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} \end{bmatrix}.$$

Thus,
$$B = \begin{array}{c} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{3} \\ 0 & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{3} \end{bmatrix} \\ \quad q_1 \qquad q_2 \end{array} \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} \end{bmatrix}.$$

Now we'll construct the full rank QR factorization by adding an orthonormal vector $q_3 = \frac{v_3}{\|v_3\|}$ to $\hat{Q}$. Like in part (a), we need $v_3$ to satisfy

$\langle q_1, v_3 \rangle = 0 \Rightarrow \frac{1}{\sqrt{2}} v_{13} + \frac{1}{\sqrt{2}} v_{33} = 0 \Rightarrow v_{13} = -v_{33}$ ①

$\langle q_2, v_3 \rangle = 0 \Rightarrow \frac{1}{\sqrt{3}} v_{13} + \frac{1}{\sqrt{3}} v_{23} - \frac{1}{\sqrt{3}} v_{33} = 0$, ① $\Rightarrow \frac{1}{\sqrt{3}} v_{23} - \frac{2}{\sqrt{3}} v_{33} = 0$

$\|q_3\| = 1.$ $\qquad\qquad \Rightarrow v_{23} = 2v_{33}.$

Let $v_{33} = 1$. Then, $v_3 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}.$

Hence, $q_3 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}.$

Therefore our full-rank QR factorization is:

$$B = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{3} & -1/\sqrt{6} \\ 0 & 1/\sqrt{3} & 2/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{3} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} \\ 0 & 0 \end{bmatrix}.$$

6)

**6.** Let $A \in \mathbb{R}^{m\times m}$ with $m \geq n$, and let $A = \hat{Q}\hat{R}$ be a reduced QR factorization.

**(a)** Show that $A$ has rank $n$ if and only if all the diagonal entries of $\hat{R}$ are nonzero.

**(b)** Suppose $\hat{R}$ has $k$ nonzero diagonal entries for some $k$ with $0 \leq k < n$. What does this imply about the rank of $A$? Exactly $k$? At least $k$? At most $k$? Give a precise answer and prove it.

a) To show this statement, we simply need to show $\text{rank}(A) = \text{rank}(\hat{Q}\hat{R}) = \text{rank}(\hat{R})$.

This is sufficient since $\hat{R}$ is upper triangular.

Proof: Let $u_1, \dots, u_k$ be a basis for the nullspace of $\hat{R}$. Then, $\dim(\text{null}(\hat{R})) = k$, and

$$\hat{R} u_i = 0, \quad \text{for } i = 1, \dots, k.$$

$$\hat{R}u_i = 0, \quad \text{for} \quad i = 1, \dots, k.$$

$$\Rightarrow \quad \hat{Q}\hat{R}u_i = 0. \quad \text{Thus} \quad \dim(\text{null}(\hat{Q}\hat{R})) \geq k.$$

Suppose towards a contradiction that $\dim(\text{null}(\hat{Q}\hat{R})) > k$.

Then, there exists $v \neq 0$ s.th $\hat{R}v \neq 0$ and $\hat{Q}\hat{R}v = 0$. However, the columns of $\hat{Q}$ are linearly independent, i.e. $\hat{Q}y = 0 \Rightarrow y = 0$. Thus,

$\hat{Q}(\hat{R}v) = 0 \Rightarrow \hat{R}v = 0$, $\lightning$. Hence $\dim(\text{null}(\hat{Q}\hat{R})) \leq k$, implying $\dim(\text{null}(\hat{Q}\hat{R})) = k$. Then, by the rank-nullity theorem,

$$\text{rank}(A) = \text{rank}(\hat{Q}\hat{R}) = n - \dim(\text{null}(\hat{Q}\hat{R})) = n - \dim(\text{null}(\hat{R}))$$
$$= \text{rank}(\hat{R}).$$

Using this result,

$\hat{R}$ is upper-triangular, so $\text{rank}(\hat{R}) = |\{r_{ii} \mid r_{ii} \neq 0\}|$. Because $\text{rank}(A) = n = \text{rank}(\hat{R})$, $\hat{R}$ must have no diagonal zero entries.
If $\hat{R}$ has no zero diagonal entries, then $\text{rank}(\hat{R}) = n$, so $\text{rank}(A) = n$.

$\blacksquare$

**b)** Since we suppose $\hat{R}$ has $k$ nonzero diagonal entries, we know $\text{rank}(\hat{R}) = k$. As proved in part (a) above, $\text{rank}(A) = \text{rank}(\hat{R})$. Thus $\text{rank}(A)$ is exactly $k$. $\blacksquare$

**7)** **7.** Let $A \in \mathbb{R}^{m \times n}$. Determine the exact numbers of floating point additions, subtractions, multiplications, and divisions involved in computing the factorization $A = \hat{Q}\hat{R}$ in the Modified Gram-Schmidt algorithm in Chapter 3 of the lecture note.

note we don't count the square root since the problem asks for $(+,-,\times)$ operations.

do $i = 1, \dots, n$

$\quad r_{ii} = \|v_i\| \longrightarrow \sqrt{\langle v_i, v_i \rangle}$ is $2n-1$ ops, $n$ times.

$\quad q_i = v_i \cdot \dfrac{1}{r_{ii}} \longrightarrow n$ ops, $n$ times

$\quad$ do $k = i+1, \dots, n$

$\quad\quad r_{ik} = q_i^T v_k \longrightarrow 2n-1$ ops, $n(n-i)$ times

$\quad\quad v_k = v_k - r_{ik} q_i \longrightarrow 2n$ ops, $n(n-i)$ times

$\quad$ end do

end do

$\Rightarrow$ Total operations:

$$n(2n-1) + n^2 + n(n-i)(2n-1) + n(n-i)2n, \quad i=1,\dots,n$$

$$= n(2n-1) + n^2 + n(n-i)(4n-1), \quad i=1,\dots,n$$

$$= 2n^2 - n + n^2 + n(4n-1)\cdot\sum_{i=1}^{n}(n-i)$$

$$= 3n^2 - n + n(4n-1)\cdot\frac{n(n-1)}{2}$$

$$= 3n^2 - n + \frac{(4n^2-n)(n^2-n)}{2}$$

$$= 3n^2 - n + \frac{4n^4 - 4n^3 - n^3 + n^2}{2}$$

$$= 2n^4 - \frac{5}{2}n^3 + \frac{7}{2}n^2 - n \quad \text{operations.}$$

8) **8.** Determine the (i) eigenvalues, (ii) determinant, and (iii) singular values of a Householder reflector. For the eigenvalues, give a geometric argument as well as an algebraic proof.

i) From 3.93 in the lecture notes, for vector $v$ with $\|v\| = 1$, the associated Householder matrix is

$$H = I - 2vv^T \in \mathbb{R}^{m\times m}.$$ We also know $H$ is symmetric and orthogonal.

Since $v \in \mathbb{R}^m$, there exists a basis of $m-1$ vectors, say $W = \{w_1, w_2, \dots, w_{m-1}\}$ s.th. $\text{span}(W) = \text{span}(v)^\perp$.

From 3.94, if we let $w \in W$, i.e. $w \perp v$, then

$$Hw = w \Rightarrow \lambda = 1, \text{ with multiplicity } m-1.$$

If we let $w \in \text{span}(v)$, i.e. $w$ is parallel to $v$, then

$$Hw = -w \Rightarrow \lambda = -1 \text{ with multplicity } 1.$$

Hence Householder matrices have eigenvalues 1 or -1.

ii) The determinant of a matrix is the product of its eigenvalues. Hence, the determinant of a Householder matrix is $(1)^{m-1}(-1) = -1$.

iii) Since a Householder matrix $H$ is symmetric and orthogonal, the singular values are

$$\sigma_i = \sqrt{\lambda_i(H^2)} = \sqrt{\lambda_i(I)} = 1, \quad \text{for all } i=1,\dots,m.$$

9)

**9.** Gram-Schmidt Process: Let

$$v_1 = \begin{bmatrix} 1 \\ \varepsilon \\ 0 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ 0 \\ \varepsilon \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \varepsilon \end{bmatrix},$$

where $\varepsilon$ is given to satisfy that its floating point operation $1+\varepsilon^2$ becomes 1 in computer arithmetic.

(a) Apply the classical Gram-Schmidt method and show that the computed vectors <u>are not</u> numerically orthogonal, i.e., computed vectors have dot products much larger than $\varepsilon$.

(b) Apply the Modified Gram-Schmidt method and show that the computed vectors <u>are</u> numerically orthogonal, i.e., computed vectors have dot products $= \mathcal{O}(\varepsilon)$.

a) We apply Graham-Schmidt to the set of vectors $V = \{V_1, V_2, V_3\}$. Note, $\|V_1\| = \|V_2\| = \|V_3\| = 1+\varepsilon^2 = 1$ in this problem.

Set $u_1 = V_1$. Then,

$$u_2 = V_2 - \frac{\langle V_1, V_2 \rangle}{\|V_1\|^2} V_1$$

$$= \begin{bmatrix} 1 \\ 0 \\ \varepsilon \\ 0 \end{bmatrix} - \frac{1}{1+\varepsilon^2} \begin{bmatrix} 1 \\ \varepsilon \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \varepsilon \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ \varepsilon \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -\varepsilon \\ \varepsilon \\ 0 \end{bmatrix}$$

$$u_3 = V_3 - \frac{\langle V_2, V_3 \rangle}{\|V_2\|^2} V_2 - \frac{\langle V_1, V_3 \rangle}{\|V_1\|^2} V_1$$

$$= \begin{bmatrix} 1 \\ 0 \\ 0 \\ \varepsilon \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ \varepsilon \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ \varepsilon \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ -\varepsilon \\ -\varepsilon \\ \varepsilon \end{bmatrix}$$

Hence, $U = \left\{ \begin{bmatrix} 1 \\ \varepsilon \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -\varepsilon \\ \varepsilon \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -\varepsilon \\ -\varepsilon \\ \varepsilon \end{bmatrix} \right\}$ should all be orthogonal.

$$\langle u_1, u_2 \rangle = -\varepsilon^2 = 0$$

$$\langle u_1, u_3 \rangle = -1-\varepsilon^2 = -1 \quad \leftarrow \text{not orthogonal.}$$

$$\langle u_2, u_3 \rangle = 0$$

b) Now we apply modified G-S.

D) Now we apply modified G-S.

Following the algorithm from Ch. 3,

$$V = \left\{ \begin{bmatrix} 1 \\ \varepsilon \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ \varepsilon \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ \varepsilon \end{bmatrix} \right\}.$$

$$\quad v_1 \qquad v_2 \qquad v_3$$

$\underline{i=1}$:

$$r_{11} = \|v_1\| = 1+\varepsilon^2 = 1.$$

$$\Rightarrow q_1 = \frac{v_1}{r_{11}} = \begin{bmatrix} 1 \\ \varepsilon \\ 0 \\ 0 \end{bmatrix}$$

$\underline{k=2}$:

$$r_{12} = q_1^T v_2 = \begin{bmatrix} 1 & \varepsilon & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \varepsilon \\ 0 \end{bmatrix} = 1$$

$$v_2 = v_2 - r_{12} q_1$$

$$v_2 = \begin{bmatrix} 1 \\ 0 \\ \varepsilon \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ \varepsilon \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -\varepsilon \\ \varepsilon \\ 0 \end{bmatrix}$$

$\underline{k=3}$

$$r_{13} = q_1^T v_3 = \begin{bmatrix} 1 & \varepsilon & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \varepsilon \end{bmatrix} = 1.$$

$$v_3 = v_3 - r_{13} q_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \varepsilon \end{bmatrix} - \begin{bmatrix} 1 \\ \varepsilon \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -\varepsilon \\ 0 \\ \varepsilon \end{bmatrix}$$

$\underline{i=2}$

$$r_{22} = \|v_2\| = \sqrt{2\varepsilon^2} = \sqrt{2}\,\varepsilon$$

$$\Rightarrow q_2 = \frac{v_2}{r_{22}} = \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

$\underline{k=3}$

$$r_{23} = q_2^T v_3 = \begin{bmatrix} 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} 0 \\ -\varepsilon \\ 0 \\ \varepsilon \end{bmatrix} = \frac{\varepsilon}{\sqrt{2}}$$

$$v_3 = v_3 - r_{23} q_2$$

$$v_3 = \begin{bmatrix} 0 \\ -\varepsilon \end{bmatrix} \quad \varepsilon \begin{bmatrix} 0 \\ -1/\sqrt{2} \end{bmatrix} - \begin{bmatrix} 0 \\ -\varepsilon \end{bmatrix} + \begin{bmatrix} 0 \\ \varepsilon/2 \end{bmatrix} = \begin{bmatrix} 0 \\ -\varepsilon/2 \\ -\varepsilon/2 \end{bmatrix}$$

$v_3 = v_3 - r_{23} q_2$

$$v_3 = \begin{bmatrix} 0 \\ -\varepsilon \\ 0 \\ \varepsilon \end{bmatrix} - \frac{\varepsilon}{\sqrt{2}} \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -\varepsilon \\ 0 \\ \varepsilon \end{bmatrix} + \begin{bmatrix} 0 \\ \varepsilon/2 \\ -\varepsilon/2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -\varepsilon/2 \\ -\varepsilon/2 \\ \varepsilon \end{bmatrix}$$

$\underline{i = 3:}$

$$r_{33} = \|v_3\| = \sqrt{\left(\frac{\varepsilon}{2}\right)^2 + \left(\frac{\varepsilon}{2}\right)^2 + \varepsilon^2} = \sqrt{\frac{3}{2}\varepsilon^2} = \sqrt{\frac{3}{2}}\,\varepsilon$$

$$q_3 = \frac{v_3}{r_{33}} = \frac{\sqrt{2}}{\sqrt{3}\,\varepsilon} \begin{bmatrix} 0 \\ -\varepsilon/2 \\ -\varepsilon/2 \\ \varepsilon \end{bmatrix} = \begin{bmatrix} 0 \\ -1/\sqrt{6} \\ -1/\sqrt{6} \\ -\sqrt{6}/3 \end{bmatrix}$$

Orthogonality check (on order of $\mathcal{O}(\varepsilon)$):

$$\langle q_1, q_2 \rangle = \begin{bmatrix} 1 & \varepsilon & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} = \frac{-\varepsilon}{\sqrt{2}} \rightsquigarrow \mathcal{O}(\varepsilon).$$

$$\langle q_1, q_3 \rangle = \begin{bmatrix} 1 & \varepsilon & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ -1/\sqrt{6} \\ -1/\sqrt{6} \\ -\sqrt{6}/3 \end{bmatrix} = \frac{-\varepsilon}{\sqrt{6}} \rightsquigarrow \mathcal{O}(\varepsilon).$$

$$\langle q_2, q_3 \rangle = \begin{bmatrix} 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} 0 \\ -1/\sqrt{6} \\ -1/\sqrt{6} \\ -\sqrt{6}/3 \end{bmatrix} = 0.$$

Hence $q_1, q_2, q_3$ are orthogonal on the order of $\mathcal{O}(\varepsilon)$.

10)

**10.** Let $P_j$ be the $m \times m$ orthogonal projector defined by $P_j = I - Q_{j-1}Q_{j-1}^T$, where $Q_{j-1}$ is a matrix whose $k$-th column is an orthogonal vector $q_k$, $1 \le k \le j-1$. Each $q_k$ is an $n$-vector satisfying $q_k \perp q_\ell$ for $k \neq \ell$. Prove that

$$P_j = P_{\perp q_{j-1}} P_{\perp q_{j-2}} \cdots P_{\perp q_1}, \quad j = 2, 3, \ldots, m, \tag{2}$$

where $P_{\perp q_k} = I - q_k q_k^T$.

$$P_{\perp q_k} = I - q_k q_k^T$$

WTS

$$P_j = P_{\perp q_{j-1}} \cdot P_{\perp q_{j-2}} \cdot \ldots \cdot P_{\perp q_1}, \quad j = 2, \ldots, m$$

Given

$$P_j = I - Q_{j-1} Q_{j-1}^T.$$

$\Rightarrow$ WTS $\quad I - Q_{j-1}Q_{j-1}^T = (I - q_{j-1}q_{j-1}^T)(I - q_{j-2}q_{j-2}^T) \cdot \ldots \cdot (I - q_1 q_1^T)$

$$(I - q_{j-1}q_{j-1}^T)(I - q_{j-2}q_{j-2}^T)$$

$$= I^2 - q_{j-1}q_{j-1}^T - q_{j-2}q_{j-2}^T - q_{j-1}(\overset{=0}{q_{j-1}^T q_{j-2}})q_{j-2}^T - \ldots$$

$$- q_1(\overset{=0}{q_1^T q_2})q_1^T - q_1 q_1^T$$

$$= I - [\underbrace{q_{j-1}q_{j-1}^T + q_{j-2}q_{j-2}^T + \ldots + q_2 q_2^T + q_1 q_1^T}]$$

Thus it suffices to show this is $Q_{j-1}Q_{j-1}^T$.

$$Q_{j-1}Q_{j-1}^T = \left[ q_1 \mid q_2 \mid \ldots \mid q_{j-1} \right] \begin{bmatrix} q_1^T \\ \hline q_2^T \\ \hline \vdots \\ \hline q_{j-1}^T \end{bmatrix}$$

$$= q_1 q_1^T + q_2 q_2^T + \ldots + q_{j-1}q_{j-1}^T.$$

▨.

11) **11.** Let $A \in \mathbb{R}^{m \times n}$. Show that $\text{cond}(A^T A) = (\text{cond}(A))^2$.

Recall $\quad \text{cond}(A) = \dfrac{\sigma_{max}(A)}{\sigma_{min}(A)}.$

The SVD of $A \in \mathbb{R}^{m \times n}$ is

$A = U\Sigma V^T$, $\quad U, V$ orthogonal, $U \in \mathbb{R}^{m\times m}$
$V \in \mathbb{R}^{n\times n}$
$\Sigma \in \mathbb{R}^{m\times n}$.

Then, $A^T = V\Sigma U^T$

$$\Rightarrow \quad A^TA = V\Sigma^T U^T U \Sigma V^T = V\Sigma^T\Sigma V^T.$$

The product $X = \Sigma^T\Sigma$ results in a diagonal $n\times n$ matrix with $X_{ii} = \Sigma_{ii}^2$, where $\underset{\sigma_{max}(A)}{\Sigma_{11}} \geq \Sigma_{22} \geq \dots \geq \underset{\sigma_{min}(A)}{\Sigma_{nn}} \geq 0$ are the singular values of $A$.

Hence, we want to show

$$\text{cond}(V X V^T) = \text{cond}(U\Sigma V^T)^2.$$

We have

$$\text{cond}(A^TA) = \text{cond}(V X V^T) = \frac{X_{11}}{X_{nn}} = \frac{\Sigma_{11}^2}{\Sigma_{nn}^2}$$

$$(\text{cond}(A))^2 = (\text{cond}(U\Sigma V^T))^2 = \left(\frac{\Sigma_{11}}{\Sigma_{nn}}\right)^2 = \frac{\Sigma_{11}^2}{\Sigma_{nn}^2}$$

$$= \text{cond}(A^TA)$$

∎