Lecture 4

Orthonormal sets of vectors and QR factorization

★ **Orthonormal set of vectors**

$\Rightarrow$ Set of vectors $u_1, \cdots u_K \in \mathbb{R}^n$ is

- Normalized if $\|u_i\| = 1$, $i = 1, \cdots K$
- Orthogonal if $u_i \perp u_j$ for $i \neq j$
- Orthonormal if both.

$\Rightarrow$ in terms of $U = [u_1, u_2 \cdots u_K]$, orthonormal means,

$$U^T U = I_K \quad (UU^T \neq I_n)$$

$\Rightarrow$ Orthonormal vectors are independent.

$u_1, u_2, \cdots, u_K$ are independent means

If $\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_K u_K = 0$ — (1)

$\Rightarrow \alpha_1 = 0, \alpha_2 = 0 \cdots \alpha_k = 0$

multiply (1) by u_i^T

$$\alpha_i u_i^T u_i = 0 \quad \Rightarrow \boxed{\alpha_i = 0}$$

$\Rightarrow$ Hence $u_1 \cdots u_K$ is an orthonormal basis for $\text{Span}(u_1, \cdots u_K) = R(U)$

→ Suppose columns of $U = [u_1, u_2 \cdots u_k]$ are orthonormal

if $\omega = Uz$, then $\|\omega\| = \|z\|$

↳ mapping $\omega = Uz$ is isometric

$z \in \mathbb{R}^{k\times 1}$, $\omega \in \mathbb{R}^{n\times 1}$, $U \in \mathbb{R}^{n\times K}$

↳ it preserves distance

$$\omega = Uz$$

$$\|\omega\|^2 = \|Uz\|^2$$

$$\Rightarrow \|\omega\|^2 = (Uz)^T(Uz) = z^T U^T U z = z^T z = \|z\|^2$$

$$\Rightarrow \boxed{\|\omega\| = \|z\|}$$

→ Inner products are also preserved

$$\langle Uz, U\tilde{z}\rangle = \langle z, \tilde{z}\rangle$$

→ norms and inner products preserved, so angles are preserved:

$$\angle(Uz, U\tilde{z}) = \angle(z, \tilde{z})$$

★ Orthonormal basis for $\mathbb{R}^n$

→ Suppose $u_1, \cdots, u_n$ is an orthonormal basis for $\mathbb{R}^n$.

→ So here $U^T U = I \Rightarrow U^T = U^{-1}$

also $UU^T = I \Rightarrow \sum_{i=1}^{n} u_i u_i^T = I$

→ $u_i u_i^T$ is called Outer product or dyad

⇒ Suppose columns of $U = [u_1, u_2 \cdots u_k]$ are orthonormal

if $w = Uz$, then $\|w\| = \|z\|$

↳ mapping $w = Uz$ is isometric ($w \in \mathbb{R}^{n\times 1}$, $z \in \mathbb{R}^{k\times 1}$, $U \in \mathbb{R}^{n\times k}$)

↳ It preserves distance

$$w = Uz$$

$$\|w\|^2 = \|Uz\|^2$$

$$\Rightarrow \|w\|^2 = (Uz)^T(Uz) = z^T U^T U z = z^T z = \|z\|^2$$

$$\Rightarrow \boxed{\|w\| = \|z\|}$$

⇒ Inner products are also preserved

$$\langle Uz, U\tilde{z}\rangle = \langle z, \tilde{z}\rangle$$

⇒ norms and inner products preserved, so angles are preserved:

$$\angle(Uz, U\tilde{z}) = \angle(z, \tilde{z})$$

★ <u>Orthonormal basis for $\mathbb{R}^n$</u>

⇒ Suppose $u_1, \cdots, u_n$ is an orthonormal basis for $\mathbb{R}^n$.

⇒ So here $U^T U = I \Rightarrow U^T = U^{-1}$

also $UU^T = I \Rightarrow \sum_{i=1}^{n} u_i u_i^T = I$

⇒ $u_i u_i^T$ is called Outer product. or dyad

★ Expansion in Orthonormal basis

⇒ Suppose U is Orthogonal, so $x = UU^T x$

$$x = \sum_{i=1}^{n} (u_i^T x) u_i$$

⇒ $u_i^T x$ is Component of x in the direction u_i

⇒ $a = U^T x = \begin{bmatrix} u_1^T x \\ u_2^T x \\ \vdots \\ u_n^T x \end{bmatrix}$ resolves x into the vector of its u_i Components.

⇒ $x = Ua$ reconstitute x from its u_i Components.

⇒ Examples of Orthonormal matrices:
- → rotations (about some fixed axis)
- → reflections (through some plane)

★ Gram Schmidt Procedure (It is an algorithm)

⇒ Given Independent set of vectors $a_1, a_2 \cdots a_K \in R^n$
G.S procedure finds orthonormal vectors $q_1, \cdots q_K$ s.t

$$\text{Span}(a_1, \cdots a_r) = \text{Span}(q_1, \cdots q_r) \quad \forall\ r \leq K$$

⇒ rough idea of method.

→ first orthogonalize each vector wrt previous one

→ then normalize result to have norm one.

① **Step 1a:** $\tilde{q}_1 = a_1$

② **Step 1b:** $q_1 = \tilde{q}_1 / \|\tilde{q}_1\|$ (Normalization)

③ **Step 2a:** $\tilde{q}_2 = a_2 - (q_1^T a_2) q_1$ (orthogonalization)

④ **Step 2b:** $q_2 = \tilde{q}_2 / \|\tilde{q}_2\|$ (Normalization)

⑤ **Step 3a:** $\tilde{q}_3 = a_3 - (q_1^T a_3) q_1 - (q_2^T a_3) q_2$

⑥ **Step 3b:** $q_3 = \tilde{q}_3 / \|\tilde{q}_3\|$

⑦ etc...

$$a_i = (q_1^T a_i) q_1 + (q_2^T a_i) q_2 + \cdots + \|\tilde{q}_i\| q_i$$

$$= r_{1i} q_1 + r_{2i} q_2 + \cdots + r_{ii} q_i$$

QR decomposition

⇒ Above can be written in matrix form

$$A = QR$$

$$\underbrace{[a_1, a_2 \cdots a_K]}_{A} = \underbrace{[q_1 \; q_2 \cdots q_K]}_{Q} \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1K} \\ 0 & r_{22} & \cdots & r_{2K} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & r_{KK} \end{bmatrix}}_{R}$$

s.

R^n

s.t

$\Rightarrow$ $Q^T Q = I_K$ and R is upper triangular & invertible.

$\Rightarrow$ Called QR decomposition (or factorization) of A.

$\Rightarrow$ Usually computed using a variation on Gram-Schmidt procedure which is less sensitive to numerical (rounding) errors.

$\Rightarrow$ Columns of Q are orthonormal basis for $R(A)$

★ <u>General Gram-Schmidt procedure</u>

$\Rightarrow$ In basis GS we assume $a_1, \dots a_K \in R^n$ are independent.

- <u>modified algorithm</u>: When we encounter $\tilde{q}_j = 0$, skip to next vector a_{j+1} and continue.

$r = 0;$

for $i = 1, 2 \dots K$

{

$$\tilde{a} = a_i - \sum_{j=1}^{r} q_j q_j^T a_i$$

if $\tilde{a} \neq 0$

{

$r = r + 1;$

$$q_r = \frac{\tilde{a}}{\|\tilde{a}\|}$$

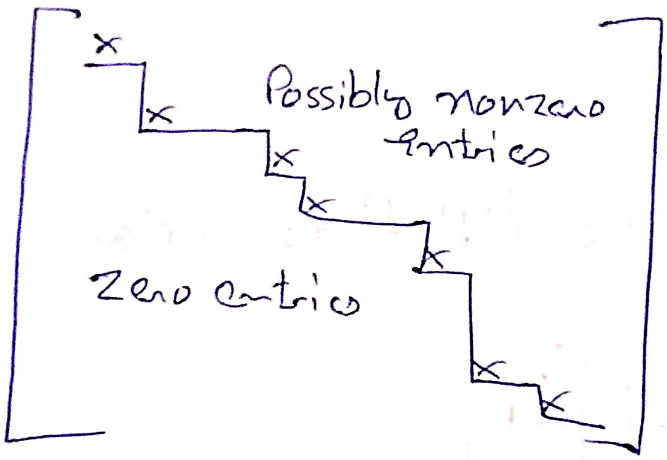
}

}

→ $q_1, \dots, q_r$ is an orthogonal basis for $R(A)$ $\{r = \text{Rank}(A)\}$

⇒ each a_i is linear combination of previously generated q_j's.

⇒ In matrix notation we have $A = QR$ with $Q^TQ = I_r$ and $R \in \mathbb{R}^{r \times k}$ in Upper Staircase form:



→ Corner entries (shown as x) are nonzero.

⇒ Can permute columns with x to front of matrix:

$\boxed{A = Q[\tilde{R}\ S]P}$ $\boxed{A = QR}$

where,

$\tilde{R} \in \mathbb{R}^{r \times r}$ is upper triangular and invertible

$R \in \mathbb{R}^{k \times k}$ is a permutation matrix

★ Applications

⇒ directly yields orthonormal basis for $R(A)$

⇒ yields factorization $A = BC$ with $B \in R^{n \times r}$ and $C \in R^{r \times k}$, $r = \text{Rank}(A)$

⇒ to check if $b \in \text{span}(a_1, \dots a_k)$: apply Gram-Schmidt to $[a_1 \dots a_k \; b]$

★ Full QR Factorization

⇒ With $A = Q_1 R_1$ the QR factorization as above write

$$A = [Q_1 \; Q_2] \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$$

where $[Q_1 Q_2]$ is orthogonal i.e. columns of $Q_2 \in R^{n \times (n-r)}$ are orthogonal, orthonormal to Q_1.

⇒ To find Q_2:

- find any matrix $\tilde{A}$ st $[A \tilde{A}]$ is full rank (e.g. $\tilde{A} = I$)
- apply general GS to $[A \tilde{A}]$
- Q_1 are orthonormal vectors obtained from columns of A.
- Q_2 are orthonormal obtained from extra column ($\tilde{A}$)

$\Rightarrow R(Q_1)$ and $R(Q_2)$ are called complementary subspaces since,

- they are orthogonal ($R(Q_1) \perp R(Q_2)$)
- there sum is $\mathbb{R}^n$ ($R(Q_1) + R(Q_2) = \mathbb{R}^n$)

(i.e: every vector in $\mathbb{R}^n$ can be expressed as a sum of two vectors, one from each subspace.)

★ Orthogonal decomposition induced by A

$$A = [Q_1 \; Q_2] \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$$

$$A^T = [R_1^T \; 0] \begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix}$$

$$A^T z = 0 = [R_1^T \; 0] \begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix} z = [R_1^T \; 0] \begin{bmatrix} Q_1^T z \\ Q_2^T z \end{bmatrix}$$

$\Rightarrow A^T z = 0$ iff $Q_1^T z = 0 \Leftrightarrow z \in R(Q_2)$

$\Rightarrow$ so $R(Q_2) = N(A^T)$ {z is orthogonal to all the columns of Q_1}

$\Rightarrow$ We conclude $R(A)$ and $N(A^T)$ are complementary subspaces.

$\Rightarrow$ Called orthogonal decomposition (of $\mathbb{R}^n$) induced by $A \in \mathbb{R}^{n \times k}$.

$$N(A) \overset{\perp}{+} R(A^T) = \mathbb{R}^k$$ {obtained by switching A to A^T}

$$\boxed{N(A^T) \overset{\perp}{+} R(A) = \mathbb{R}^n}$$

(four fundamental subspaces)