Lecture XVIII : §3.6 Orthogonal basis for Subspaces

TODAY: Define inner products (generalizing dot product in $\mathbb{R}^n$) & find basis that are well-behaved with respect to this inner product.

§1. Inner products of subspaces $\mathbb{W}$ of $\mathbb{R}^n$

Def: An inner product for (a vector space) $\mathbb{W}$ is a function

$$\langle -, - \rangle : V \times V \longrightarrow \mathbb{R}$$

assigning $(u, v) \longmapsto \langle u, v \rangle$ in $\mathbb{R}$ for u, v in $\mathbb{W}$, and satisfying for all u, v, w in $\mathbb{W}$:

(1) $\langle u, u \rangle \geq 0$, $\langle u, u \rangle = 0$ if and only if $u = \mathbb{0}$ ("non-degen. condition")

(2) $\langle u, v \rangle = \langle v, u \rangle$ [Symmetric]

(3) $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle = \langle u, \alpha v \rangle$ for any scalar α.

(4) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$

Note: By symmetry, we ~~also~~ have:

(4') $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$.

(2) – (4') bilinear form.

Examples: ① $\langle u, v \rangle = u^T \cdot v$ for u, v in $\mathbb{R}^n$
$= u \bullet v$ (usual dot product)

(1) becomes $\|u\|^2 = 0$ & $\|u\| = 0$ if & only if $u = \mathbb{0}$.

(2) – (4) follow from matrix multiplication.

② Q symmetric matrix of rank n. Then $\langle u, v \rangle := u^T Q v$.

(2) $\langle v, u \rangle = v^T Q u \underset{\substack{\downarrow \\ \text{in } \mathbb{R}}}{=} (v^T Q u)^T = u^T \; Q^T (v^T)^T = u^T Q v = \langle u, v \rangle$.

(3) – (4) True

(1) not always True. Eg $Q = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ $\mathbb{W} = \mathbb{R}^n$ $[u_1, u_2] \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = u_1^2 - u_2^2$

• If $u_1 = u_2 = 1$ we get $\langle u, u \rangle = 0$ but $\begin{bmatrix} 1 \\ 1 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Obs: Condition that guarantees (1) to hold? A: Q "positive definite" ["all eigenvalues are > 0"]

Obs: An inner product defines a norm on $\mathbb{W}$: $\| v \| = \sqrt{\underbrace{\langle v, v \rangle}_{\geq 0}}$.

§2. Orthogonal bases for $\mathbb{R}^n$

From now on, we work with the classical inner product in $\mathbb{R}^n$:

Def If u, v are vectors in $\mathbb{R}^n$, we say u & v are orthogonal of perpendicular if $\langle u,v\rangle = u^T\cdot v = 0$. Write $u \perp v$.

Def: A set of vectors $S=\{v_1,\dots,v_p\}$ in $\mathbb{R}^n$ is orthogonal if each pair v_i, v_j with $i\neq j$ is orthogonal

Example: $\{\underbrace{\begin{bmatrix}1\\0\\0\end{bmatrix}}_{=v_1}, \underbrace{\begin{bmatrix}0\\1\\-1\end{bmatrix}}_{=v_2}, \underbrace{\begin{bmatrix}0\\1\\1\end{bmatrix}}_{=v_3}\}$ is an orthogonal set in $\mathbb{R}^n$ because $v_1^Tv_2 = v_1^Tv_3 = v_2^Tv_3 = 0$.

- Why do we care? These sets are linearly independent unless they contain $\mathbb{0}$.

Theorem: If S is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is l.i.

Proof: Write $S=\{\vec{v}_1,\dots,\vec{v}_p\}$ & $a_1\vec{v}_1+\dots+a_p\vec{v}_p = \begin{bmatrix}0\\\vdots\\0\end{bmatrix}$

Then $0 = \vec{v}_1^T\cdot\begin{bmatrix}0\\\vdots\\0\end{bmatrix} = v_1^T(a_1\vec{v}_1+\dots+a_p\vec{v}_p) = a_1 v_1^Tv_1 + a_2\overbrace{v_1^Tv_2}^{=0}+\dots$
$\dots + a_p\underbrace{v_1^Tv_p}_{=0}$
$= a_1\|\vec{v}_1\|^2$

Since $\vec{v}_1\neq\mathbb{0}$, then $\|\vec{v}_1\|\neq 0$ So $0 = a_1\underbrace{\|\vec{v}_1\|^2}_{\neq 0}$ forces $a_1=0$

Similarly: $0 = \vec{v}_2^T\begin{bmatrix}0\\\vdots\\0\end{bmatrix} = a_1\underbrace{v_2^Tv_1}_{=0} + a_2v_2^Tv_2+\dots+a_p\underbrace{v_2^Tv_p}_{=0} = a_2\|v_2\|^2$

$\vdots$

$0 = \vec{v}_p^T\begin{bmatrix}0\\\vdots\\0\end{bmatrix} = a_1\underbrace{v_p^Tv_1}_{=0} + a_2\underbrace{v_p^Tv_2}_{=0}+\dots+a_pv_p^Tv_p = a_p\|v_p\|^2$

Since $\vec{v}_2,\dots,\vec{v}_p\neq\mathbb{0}$, we conclude $a_1=a_2=\dots=a_p=0$ & so S is l.i. $\square$

Def: Fix $W\neq\{0\}$ subspace of $\mathbb{R}^n$ with basis $B=\{w_1,\dots,w_p\}$

- We say B is an orthogonal basis if B is a basis consisting of orthogonal vectors.
- ________ orthonormal basis if B is an orthogonal basis & $\|w_1\|=\|w_2\|=\dots=\|w_p\|=1$ (all vectors have norm 1).

Example: $W = \mathrm{Span}\left(\underbrace{\begin{bmatrix}1\\-1\\1\end{bmatrix}}_{=v_1}, \underbrace{\begin{bmatrix}1\\2\\1\end{bmatrix}}_{v_2}\right\}$ subspace of $\mathbb{R}^3$

- $B=\{v_1,v_2\}$ is orthogonal basis : $v_1^Tv_2 = 1-2+1 = 0$. (so l.i.)

- Not orthonormal basis $\left\| \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\| = \sqrt{1+1+1} = \sqrt{3}$, $\left\| \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\| = \sqrt{1+4+1} = \sqrt{6}$.

$\rightsquigarrow$ We can turn B into an orthonormal basis as follows:

$$B' = \left\{ \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\}.$$

In general: $B = \{w_1, \ldots w_p\}$ orthogonal vectors $\rightsquigarrow B' = \left\{ \frac{w_1}{\|w_1\|}, \ldots, \frac{w_p}{\|w_p\|} \right\}$ orthonormal basis to $W = Sp(w_1, \ldots, w_p)$.

- Advantage of orthonormal basis: Computing coordinates is easy!

Theorem: Fix $B = \{v_1, \ldots v_p\}$ is an orthonormal basis for $W \neq \{0\}$ subspace of $\mathbb{R}^n$.

Then, for each v in W: $[v]_B = \begin{bmatrix} v^T \cdot v_1 \\ \vdots \\ v^T v_p \end{bmatrix}$.

Proof: By definition $[v]_B = \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix}$ means $v = a_1 \vec{v}_1 + \cdots + a_p \vec{v}_p$.

Then $v^T \cdot v_1 = (a_1 v_1 + \cdots + a_p v_p)^T v_1 = a_1 \underbrace{v_1^T v_1}_{=1} + a_2 \underbrace{v_2^T v_1}_{=0} + \cdots + a_p \underbrace{v_p^T v_1}_{=0}$

So $v^T \cdot v_1 = a_1$

Similarly for the others: $v^T v_i = a_1 \cdot 0 + \cdots + a_{i-1} 0 + a_i \underbrace{v_i^T v_i}_{=1} + 0 + \cdots + 0 = a_i$.

Note: If B is an orthogonal basis $B = \{v_1, \ldots, v_p\}$ then:

$$[v]_B = \begin{bmatrix} \frac{v^T v_1}{\|v_1\|^2} \\ \vdots \\ \frac{v^T v_p}{\|v_p\|^2} \end{bmatrix}$$

$v = a_1 \frac{v_1}{\|v_1\|} + \cdots + a_p \frac{v_p}{\|v_p\|}$

for $a_1 = v^T \frac{v_1}{\|v_1\|}$, $a_2 = v^T \frac{v_2}{\|v_2\|}$, $\ldots$, $a_p = v^T \frac{v_p}{\|v_p\|}$

So $v = \boxed{\frac{v^T v_1}{\|v_1\|^2}} v_1 + \boxed{\frac{v^T v_2}{\|v_2\|^2}} v_2 + \cdots + \boxed{\frac{v^T v_p}{\|v_p\|^2}} v_p$

§3. Gram-Schmidt Algorithm:

- Input: $B = \{\vec{w}_1, \ldots, \vec{w}_p\}$ basis for a subspace W of $\mathbb{R}^n$.
- Output: An orthogonal basis for W $= \{\vec{u}_1, \ldots, \vec{u}_p\}$.
- Routine:

$$u_1 = w_1$$

$$u_2 = w_2 - \frac{u_1^T w_2}{\|u_1\|^2} u_1$$

$$u_3 = w_3 - \frac{u_1^T w_3}{\|u_1\|^2} u_1 - \frac{u_2^T w_3}{\|u_2\|^2} u_2$$

$$\vdots$$

In general: for $j=2,\dots,p$

$$u_j = w_j - \boxed{\frac{u_1^T w_j}{\|u_1\|^2}} u_1 - \boxed{\frac{u_2^T w_j}{\|u_2\|^2}} u_2 - \cdots - \boxed{\frac{u_{j-1}^T w_j}{\|u_{j-1}\|^2}} u_{j-1}.$$

scalars

Q: Why these scalars?

- Assume $\{u_1, \dots u_{j-1}\}$ is an orthogonal set.
- Need to check: $u_1^T u_j = u_2^T u_j = \cdots = u_{j-1}^T u_j = 0$. Write $u_j = w_j - a_1 u_1 - \cdots - a_{j-1} u_{j-1}$ T.B.D. ($\{u_1, \dots u_j\}$ will be l.i.)

$$u_1^T u_j = u_1^T (w_j - a_1 u_1 - a_2 u_2 - \cdots - a_{j-1} u_{j-1}).$$

$$= u_1^T w_j - a_1 \underbrace{u_1^T u_1}_{\|u_1\|^2} - a_2 \underbrace{u_1^T u_2}_{=0} - \cdots - a_{j-1} \underbrace{u_1^T u_{j-1}}_{=0}$$

So to get $0 = u_1^T u_j$ we must satisfy: $0 = u_1^T w_j - a_1 \|u_1\|^2$

$$\rightsquigarrow \boxed{a_1 = \frac{u_1^T w_j}{\|u_1\|^2}}.$$

Similarly $i = 1, \dots, j-1$

$$0 \stackrel{?}{=} u_i^T u_j = u_i^T (w_j - a_1 u_1 - \cdots - a_i u_i - \cdots - a_{j-1} u_{j-1})$$

$$0 = u_i^T w_j - a_1 \underbrace{u_i^T u_1}_{=0} - \cdots - a_i \underbrace{u_i^T u_i}_{=\|u_i\|^2} - \cdots - a_{j-1} \underbrace{u_i^T u_{j-1}}_{=0}$$

$$0 = u_i^T w_j - a_i \|u_i\|^2$$ if & only if $\boxed{a_i = \frac{u_i^T w_j}{\|u_i\|^2}}$.

- By construction $Sp(u_1, \dots u_j) = Sp(w_1, \dots w_j)$ for all $j = 1, \dots, p$ so in the end, we get $Sp(u_1, \dots u_p) = \mathbb{W}$ & $\dim \mathbb{W} = p$ so $B = \{u_1, \dots u_p\}$ is a basis. B is an orthogonal set by construction.

Example: $\mathbb{W} = \mathbb{R}^3$ with basis $\left\{ \underset{w_1}{\begin{bmatrix}1\\0\\0\end{bmatrix}}, \underset{w_2}{\begin{bmatrix}1\\2\\0\end{bmatrix}}, \underset{w_3}{\begin{bmatrix}1\\-3\\4\end{bmatrix}} \right\} \xrightarrow{\text{G.S.}} \left\{ \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\2\\0\end{bmatrix}, \begin{bmatrix}0\\0\\4\end{bmatrix} \right\}$

- $u_1 = \begin{bmatrix}1\\0\\0\end{bmatrix} \rightsquigarrow u_1^T w_2 = 1, \; u_1^T w_3 = 1, \; \|u_1\|^2 = 1$
- $u_2 = w_2 - \frac{u_1^T w_2}{\|u_1\|^2} u_1 = \begin{bmatrix}1\\2\\0\end{bmatrix} - \frac{1}{1}\begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}0\\2\\0\end{bmatrix}$

$\rightsquigarrow u_1^T w_3 = 1, \quad u_2^T w_3 = -6 \quad \|u_2\|^2 = 4$

$$u_3 = w_3 - \frac{u_1^T w_3}{\|u_1\|^2} u_1 - \frac{u_2^T w_3}{\|u_2\|^2} u_2 = \begin{bmatrix}1\\-3\\4\end{bmatrix} - \frac{1}{1}\begin{bmatrix}1\\0\\0\end{bmatrix} - \left(\frac{-6}{4}\right)\begin{bmatrix}0\\2\\0\end{bmatrix} = \begin{bmatrix}1\\-3\\4\end{bmatrix} - \begin{bmatrix}1\\0\\0\end{bmatrix} + \begin{bmatrix}0\\3\\0\end{bmatrix} = \begin{bmatrix}0\\0\\4\end{bmatrix}$$