Lecture XX: §3.6 Orthogonal bases for Subspaces

Recall: Given an (abstract) vector space $\mathbb{V}$, a ~~set~~ of vectors $B = \{v_1, \ldots, v_p\}$ is a <u>basis</u> for $\mathbb{V}$ if

(1) B spans $\mathbb{V}$ (every vector is a linear comb of $v_1, \ldots, v_p$)

(2) B is linearly independent (only solution to $\underline{\alpha_1} v_1 + \cdots + \underline{\alpha_p} v_p = \mathbb{0}$ is the trivial one: $\alpha_1 = \cdots = \alpha_p = 0$) zero vector in $\mathbb{V}$

Using basis: identify $\mathbb{V}$ with $\mathbb{R}^p$ via $v \longleftrightarrow [v]_B = \begin{array}{c} \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} \end{array} \begin{array}{c} v_1 \\ \vdots \\ v_p \end{array}$

Facts:
- Coordinates with respect to a basis respect $+, \cdot$. $\quad v = \underbrace{a_1 v_1 + \cdots + a_p v_p}_{\text{scalars}}$ (vectors)
  & $[\mathbb{0}]_B = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$
- l.i & spanning can be detected by $[\;]_B$.

Corollary:
- All bases of $\mathbb{V}$ have the same number of vectors, which we define as the <u>dimension</u> of $\mathbb{V}$
- If $\dim \mathbb{V} = n$ & $\mathbb{W}$ is a subspace of $\mathbb{V}$ of dimension n, then $\mathbb{W} = \mathbb{V}$

Thm: Assume $\mathbb{V}$ is a vector space of dimension n, and take $\{v_1, \ldots, v_n\}$ vectors in $\mathbb{V}$

(1) If $\{v_1, \ldots, v_n\}$ is l.i, then it is a basis.

(2) If $\{v_1, \ldots, v_n\}$ spans $\mathbb{V}$, _______________

(3) Any set of $n+1$ or more vectors in $\mathbb{V}$ is l.d. ; any set of less than n vectors can't span $\mathbb{V}$.

TODAY: Define inner products (generalizing dot product in $\mathbb{R}^n$) & find bases behaving well with respect to this inner product.

§1 Inner products

Definition: An <u>inner product</u> for a vector space $\mathbb{V}$ is a function $\langle -,- \rangle$ assigning a real number (call it $\langle u, v \rangle$) to a pair of vectors u, v in $\mathbb{V}$, satisfying:

(1) $\langle u, v \rangle \geq 0$, $\quad \langle u, u \rangle = 0$ if and only if $u = \mathbb{0}$ (zero vector in $\mathbb{V}$)

(2) $\langle u, v \rangle = \langle v, u \rangle \quad$ (symmetric)

(3) $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$ for any scalar α in $\mathbb{R}$

(4) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ for any u, v, w in $\mathbb{V}$

Note: Using (2), (3) & (4) also hold for the second leg.

(3') $\langle u, \alpha v \rangle = \alpha \langle u, v \rangle$

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

Examples: ① $\langle u, v \rangle = u^T v$ for u, v in $\mathbb{R}^n$ defines an inner product for $\mathbb{R}^n$
$(= u \cdot v$ usual dot product for $\mathbb{R}^2$ & $\mathbb{R}^3)$

② A symmetric 2×2 matrix $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ with $a>0$, $ad-bc>0$ (Example $A=\begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$) Then $\langle u,v\rangle = u^T A v$. is an inner product for $\mathbb{R}^2$

③ In $\mathbb{P}_2$: $\langle p,q\rangle = p(0)q(0) + p(1)q(1) + p(2)q(2)$. defines an inner product for $\mathbb{P}_2$

④ In $C_{[0,1]}$: $\langle f(x), g(x)\rangle = \int_0^1 f(x)\, g(x)\, dx$ is an inner product.

Using $\langle\, ,\rangle$ we can define the norm (or magnitud) of vectors in $\mathbb{V}$.

Def: $\|v\| = \sqrt{\langle v,v\rangle}$. (Note: $\langle v,v\rangle \geq 0$ for all v in $\mathbb{V}$)

§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 (or perpendicular) if $u^T \cdot v = 0$.

Def A set of vectors $S=\{v_1,\dots,v_p\}$ in $\mathbb{R}^n$ is orthogonal if each pair of vectors in S is orthogonal, that is $v_i^T v_j = 0$ for all $i\neq j$.

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

Why? These sets are ALWAYS l.i unless they have the zero vector in them.

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

Proof: Write $S=\{v_1,\dots,v_p\}$ & $a_1 v_1 + \dots + a_p v_p = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$

Then $0 = v_1^T \cdot \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = a_1 v_1^T v_1 + a_2 \underbrace{v_1^T v_2}_{=0} + \dots + a_p \underbrace{v_1^T v_p}_{=0} = a_1 \|v_1\|^2$

Since $v_1 \neq 0$, then $\|v_1\| \neq 0$, so from $0 = a_1 \|v\|^2$, we conclude $\boxed{a_1 = 0}$

Similarly $0 = v_2^T \cdot \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = a_2 \|v_2\|^2$ implies $a_2 = 0$ (because $v_2 \neq 0$)

$\vdots$

$0 = v_p^T \cdot \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = a_p \|v_p\|^2$ —— $a_p = 0$ (—— $v_p \neq 0$)

We conclude $a_1 = a_2 = \dots = a_p = 0$ so S is l.i. □

Definition Let $\mathbb{W}$ be a subspace of $\mathbb{R}^n$ with basis $B = \{w_1, \cdots, w_p\}$

- We say B is an orthogonal basis if B is an orthogonal set (& a basis)
- ———— orthonormal basis if B is an orthogonal basis & $\|w_1\| = \cdots = \|w_p\| = 1$ (all vectors have norm $= 1$)

Example: The plane $\mathbb{W}$ spanned by $B = \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\}$ in $\mathbb{R}^3$

- B is orthogonal basis $[1 \; -1 \; 1] \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = 1 - 2 + 1 = 0.$
- Not orthonormal basis $\left\| \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\|^2 = \sqrt{1+1+1} = \sqrt{3}$, $\left\| \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\| = \sqrt{1+4+1} = \sqrt{6}$.

- How to make it orthonormal?

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

In general: replace each vector v in B by $\frac{v}{\|v\|}$ if B is orthogonal

Why? Determining coordinates with respect to an orthonormal basis is EASY!

Thm $B = \{v_1, \ldots, v_p\}$ be a orthonormal basis for $\mathbb{W}$ (subspace of $\mathbb{R}^n$) & given v in $\mathbb{W}$

$$[v]_B = \begin{array}{c} \begin{bmatrix} v^T \cdot v_1 \\ \vdots \\ v^T \cdot v_p \end{bmatrix} \end{array} \begin{array}{c} v_1 \\ \vdots \\ v_p \end{array}$$

Proof $v = a_1 v_1 + \cdots + a_p v_p$, then $v^T \cdot v_1 = a_1 \underbrace{v_1^T v_1}_{=1} + a_2 \underbrace{v_2^T v_1}_{=0} + \cdots + a_p \underbrace{v_p^T v_1}_{=0}$

So $a_1 = v^T v_1$ $\quad = a_1$ orthonormal (orthogonal basis)

Similarly for the others : $a_i = v^T \cdot v_i$ for $i = 1, \ldots, n$.

Note: If B is an orthogonal basis, then $[v]_B = \begin{bmatrix} \frac{v^T v_1}{\|v_1\|^2} \\ \vdots \\ \frac{v^T v_p}{\|v_p\|^2} \end{bmatrix}$

§3 Gram-Schmidt Algorithm

- Input: A basis $\{w_1, \ldots w_p\}$ for a subspace $\mathbb{W}$ of $\mathbb{R}^n$.
- Output: An orthogonal basis for $\mathbb{W}$ $= \{u_1, \ldots, u_p\}$
- Routine:

$$u_1 = w_1$$

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

$$u_3 = w_3 - \underbrace{\frac{u_1^T w_3}{\|u_1\|^2}}_{\text{scalars}} u_1 - \underbrace{\frac{u_2^T w_2}{\|u_2\|^2}}_{\text{scalars}} u_2$$

$$\vdots$$

In general: $u_{j+1} = w_{j+1} - \boxed{\dfrac{u_1^T w_{j+1}}{\|u_1\|^2}}\, u_1 - \cdots - \boxed{\dfrac{u_j^T w_{j+1}}{\|u_j\|^2}}\, u_j$
(for $j = 1, \ldots, p-1$)

scalars

Q: Why these scalars?

$$u_{j+1} = w_{j+1} - a_1 u_1 - \cdots - a_j u_j \quad \& \quad \{u_1, \ldots, u_j\} \text{ orthogonal set.}$$

$$\underset{\text{want}}{0} = u_1^T u_{j+1} = u_1^T w_{j+1} - a_1 \|u_1\|^2 - a_2 \underbrace{u_1^T u_2}_{=0} - \cdots - a_j \underbrace{u_1^T u_j}_{=0}$$

$$= u_1^T w_{j+1} - a_1 \|u_1\|^2 \quad \text{so} \quad a_1 = \frac{u_1^T w_{j+1}}{\|u_1\|^2}$$

Similarly: $0 = u_i^T u_{j+1} = u_i^T w_{j+1} - 0 - \cdots - a_i \|u_i\|^2 - 0 \cdots$

so $a_i = \dfrac{u_i^T w_{j+1}}{\|u_i\|^2}$ for all $i = 1, \ldots, j$.

Proof: • Our choice of scalars ensures the vectors are pairwise orthogonal.

Claim: $u_j \neq \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$ for all j

(1) $u_1 \neq 0$ because $w_1 \neq 0$ ✓

(2) $\mathrm{Sp}(u_1 \ldots u_j) = \underbrace{\mathrm{Sp}(w_1, \ldots, w_j)}_{\text{dimension} = j}$ for all j, so $\{u_1 \ldots u_j\}$ spans a subspace of dim $= j$ so they are all non-zero vectors

Why? • Clear for $j = 1$

• For $j > 1$: $\mathrm{Sp}(\underbrace{u_1 \ldots u_j}_{\text{same span as } w_1 \cdots w_j}\ u_{j+1}) = \mathrm{Sp}(u_1 \cdots u_j, w_{j+1}) =$ because $u_{j+1} = w_{j+1} - a_1 u_1 - \cdots - a_j u_j$

$= \mathrm{Sp}(w_1 \cdots w_j, w_{j+1})$.

Example: $W = \mathbb{R}^3$ with basis $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix} \right\} \overset{GS}{\rightsquigarrow} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix} \right\}$

• $u_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \rightsquigarrow u_1^T w_2 = 1, \quad u_1^T w_3 = 1, \quad \|u_1\| = 1$

• $u_2 = w_2 - \dfrac{u_1^T w_2}{\|u_1\|^2} u_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} - \dfrac{1}{1} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$

$\rightsquigarrow u_2^T w_2 = 4, \quad u_2^T w_3 = -6, \quad \|u_2^T\| = 4$

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