$\Rightarrow$ columns of P are eigenvectors of A.

$\exists$ n columns & they are l.i. because P is invertible (why ?)

$\Rightarrow$ A has n l.i. eigenvectors.

"$\Leftarrow$" can be similarly proved

A Procedure for Diagonalizing an $n \times n$ Matrix

Step 1. Determine first whether the matrix is actually diagonalizable by searching for n linearly independent eigenvectors. One way to do this is to find a basis for each eigenspace and count the total number of vectors obtained. If there is a total of n vectors, then the matrix is diagonalizable, and if the total is less than n, then it is not.

Step 2. If you ascertained that the matrix is diagonalizable, then form the matrix $P = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_n]$ whose column vectors are the n basis vectors you obtained in Step 1.

Step 3. $P^{-1}AP$ will be a diagonal matrix whose successive diagonal entries are the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ that correspond to the successive columns of P.

Find a matrix P that diagonalizes

$$A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$$

Steps: ① Solve $\det(A - xI) = 0$ (cubic poly.) and find roots $\lambda_1, \lambda_2, \lambda_3$.

② Take λ_1. Find basis for E_{λ_1}

$$= \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} : A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda_1 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right\}$$

Let the basis vectors be $\{v_1, v_2, \cdots, v_r\}$.

r = GM of λ_1 = max. # of l.i. eigenvectors for λ_1.

③ Repeat for λ_2 & λ_3. & find $GM(\lambda_2)$ & $GM(\lambda_3)$.

④ If $\sum_{i=1}^{3} GM(\lambda_i) = n = 3$, then A is diagonalizable, & the matrix P which diagonalizes A is got by putting these 3 l.i. eigenvectors as columns.

Sol^n: $\det(A - \lambda I) = 0$

$$\begin{vmatrix} -\lambda & 0 & -2 \\ 1 & 2-\lambda & 1 \\ 1 & 0 & 3-\lambda \end{vmatrix} = 0$$

$\Rightarrow -\lambda(2-\lambda)(3-\lambda) + 2(2-\lambda) = 0$

$\Rightarrow (2-\lambda)(\lambda^2 - 3\lambda + 2) = 0$

$\Rightarrow \lambda = 2$ or $\lambda = 1$

$AM(2) = 2$ $\qquad$ $AM(1) = 1$.

THEOREM 5.2.4 Geometric and Algebraic Multiplicity

If A is a square matrix, then: $1 \leq GM \leq AM$

(a) *For every eigenvalue of A, the geometric multiplicity is less than or equal to the algebraic multiplicity.*

(b) *A is diagonalizable if and only if the geometric multiplicity of every eigenvalue is equal to the algebraic multiplicity.*

$$\left.\begin{aligned} AM(\lambda_1) &= GM(\lambda_1) = 1 \\ AM(\lambda_2) &= GM(\lambda_2) = 2 \end{aligned}\right\} \text{check!}$$

$\Rightarrow$ A is diagonalizable

with $P = \begin{bmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ } check!

Find P^{-1}:

Aug. mat. $\left[\begin{array}{ccc|ccc} -2 & -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$

$R_1 \leftrightarrow R_2$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ -2 & -1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

$R_2 \to R_2 + 2R_1$, $R_3 \to R_3 - R_1$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 1 & 2 & 0 \\ 0 & 1 & -1 & 0 & -1 & 1 \end{array}\right]$$

$R_2 \to -R_2$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & -1 & -2 & 0 \\ 0 & 1 & -1 & 0 & -1 & 1 \end{array}\right]$$

$R_3 \to R_3 - R_2$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & -1 & -2 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right]$$

$R_1 \to R_1 - R_3, \quad R_2 \to R_2 + 2R_3$

$$\left[\underbrace{\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}}_{I}\;\middle|\;\underbrace{\begin{array}{ccc} -1 & 0 & -1 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \end{array}}_{P^{-1}}\right]$$

Hence $D = P^{-1}AP$

$$= \begin{bmatrix} -1 & 0 & -1 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & 0 & -1 \\ 2 & 0 & 4 \\ 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

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

Section 6.1 – Inner products

Let V be a vector space (over $\mathbb{R}$)
We have studied $+ : V \times V \to V$
$\cdot : \mathbb{R} \times V \to V$

Is there an operation possible from $V \times V \to \mathbb{R}$?

Eg. If $V = \mathbb{R}^n$, we have a dot product
$u \cdot v = u_1 v_1 + \cdots + u_n v_n$.

For a general vector space V, we will call such an operation as an "<u>inner product</u>".

<u>Def</u>. An inner product on V is a function from $V \times V \to \mathbb{R}$, $(u, v) \to \langle u, v \rangle$ s.t. $\forall\, u, v, w \in V$ & $\forall\, k \in \mathbb{R}$,

1. $\langle u, v \rangle = \langle v, u \rangle$ (symmetry)
2. $\langle u+v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ (additivity)
3. $\langle ku, v \rangle = k \langle u, v \rangle$ (Homogeneity)
4. $\langle v, v \rangle \geq 0$ and $= 0$ iff $v = 0_v$ (positivity)
   (zero vector)

If V has such a function defined on it, then it is called an <u>inner product space</u>.

Eg. 1. $\mathbb{R}^n$ with $\langle x, y \rangle = x \cdot y$ (usual dot product) eg. ② For $\mathbb{R}^2$ with $\langle u, v\rangle = \frac{1}{9}u_1v_1 + \frac{1}{4}u_2v_2$, show that the inner product axioms hold. (An example of a 'weighted' dot product.)

Ex. 1 If $V = M_{nn}$, then $\langle A, B\rangle = tr(A^TB)$ is an inner product space.

- $tr(X^T) = tr(X)$
- $tr(X+Y) = tr(X) + tr(Y)$
- $tr(cX) = c\,tr(X)$

i) $\langle A, B\rangle = \langle B, A\rangle$

$$\begin{aligned}\langle A, B\rangle &= tr(A^TB)\\ &= tr((A^TB)^T)\\ &= tr(B^T(A^T)^T)\\ &= tr(B^TA)\\ &= \langle B, A\rangle.\end{aligned}$$

(ii) $$\begin{aligned}\langle A+C, B\rangle &= tr((A+C)^TB)\\ &= tr((A^T+C^T)B)\\ &= tr(A^TB + C^TB)\\ &= tr(A^TB) + tr(C^TB)\\ &= \langle A, B\rangle + \langle C, B\rangle.\end{aligned}$$

(iii) $$\begin{aligned}\langle kA, B\rangle &= tr((kA)^TB)\\ &= tr(kA^TB)\\ &= k\,tr(A^TB)\\ &= k\langle A, B\rangle.\end{aligned}$$

(iv) $\langle A, A \rangle$

Let $A = (A_{ij}) \in M_{nn}$.

Then

$$\langle A, A \rangle = tr(A^T A)$$

$$= \sum_{i=1}^{n} (A^T A)_{ii} = \sum_{i=1}^{n} \sum_{j=1}^{n} A^T_{ij} A_{ji}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ji}^2$$

$$\geq 0$$

$\langle A, A \rangle = 0$

$$\iff \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij}^2 = 0$$

$$\iff A_{ij} = 0 \quad \forall\, i, j, \quad 1 \leq i \leq n, \ 1 \leq j \leq n$$

$$\iff A = 0.$$ (matrix consisting of all zero entries)

$\Rightarrow \langle A, B \rangle := tr(A^T B)$ is an inner product on M_{nn}.

Note: $\|A\| = \sqrt{\langle A, A \rangle} = \sqrt{tr(A^T A)}$.

norm

Norm and distance:

DEFINITION 2 If V is a real inner product space, then the ***norm*** (or ***length***) of a vector $\mathbf{v}$ in V is denoted by $\|\mathbf{v}\|$ and is defined by

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

and the ***distance*** between two vectors is denoted by $d(\mathbf{u}, \mathbf{v})$ and is defined by

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle}$$

A vector of norm 1 is called a ***unit vector***.

Eg. In $\mathbb{R}^3$, with the usual dot product,

$$\|v\| = \sqrt{\langle v, v\rangle} = \sqrt{v \cdot v} = \sqrt{v_1^2 + v_2^2 + v_3^2}$$ } length of a vector in $\mathbb{R}^3$

$$\|u - v\| = \sqrt{\langle u-v, u-v \rangle}$$

$$= \sqrt{(u-v)\cdot(u-v)} = \underbrace{\sqrt{(u_1-v_1)^2 + (u_2-v_2)^2 + (u_3-v_3)^2}}_{\text{dist. between vectors } u \,\&\, v \text{ in } \mathbb{R}^3.}$$

THEOREM 6.1.1 *If* $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in a real inner product space* V, *and if* k *is a scalar, then:*

(*a*) $\|\mathbf{v}\| \geq 0$ *with equality if and only if* $\mathbf{v} = \mathbf{0}$.

(*b*) $\|k\mathbf{v}\| = |k|\|\mathbf{v}\|$.

(*c*) $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$.

(*d*) $d(\mathbf{u}, \mathbf{v}) \geq 0$ *with equality if and only if* $\mathbf{u} = \mathbf{v}$.

Properties of norm and distance ↑

Properties of inner product

THEOREM 6.1.2 *If* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in a real inner product space* V, *and if* k *is a scalar, then:*

(*a*) $\langle \mathbf{0}, \mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{0}\rangle = 0$

(*b*) $\langle \mathbf{u}, \mathbf{v}+\mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{u}, \mathbf{w}\rangle$

(*c*) $\langle \mathbf{u}, \mathbf{v}-\mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle - \langle \mathbf{u}, \mathbf{w}\rangle$

(*d*) $\langle \mathbf{u}-\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{w}\rangle - \langle \mathbf{v}, \mathbf{w}\rangle$

(*e*) $k\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}, k\mathbf{v}\rangle$ $= \langle ku, v\rangle$

THEOREM 6.2.1 Cauchy–Schwarz Inequality

If $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in a real inner product space* V, *then*

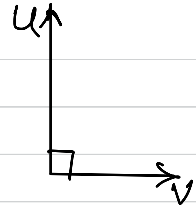
$$|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}\|\,\|\mathbf{v}\|$$

THEOREM 6.2.2 *If* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in a real inner product space* V, *and if* k *is any scalar, then:*

(*a*) $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ **[Triangle inequality for vectors]**

(*b*) $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$ **[Triangle inequality for distances]**

Sect. 6.2 — Orthogonality



Let V be an inner product space.

Def. Two vectors $u, v \in V$ are said to be orthogonal if $\langle u, v \rangle = 0$.

Ex. 2 Take $\mathbb{R}^2$ with the inner product

1. $\langle x, y \rangle_1 = x_1 y_1 + x_2 y_2$
2. $\langle x, y \rangle_2 = 3x_1 y_1 + 2x_2 y_2$

Check if $u = (1, 1)$ and $v = (1, -1)$ are orthogonal w.r.t. 1 orthogonal
— w.r.t. 2 not orthogonal

Def. Let $S = \{v_1, v_2, \dots, v_n\} \subseteq V$.

- We say S is an orthogonal (OG) set if $\langle v_i, v_j \rangle = 0 \quad \forall\, i \neq j$.

- We say S is an orthonormal (ON) set if $\langle v_i, v_j \rangle = 0 \quad \forall\, i \neq j$
  & $\langle v_i, v_i \rangle = 1 \quad \forall\, i$ (that is, $\|v_i\|^2 = 1 \ \forall\, i$)

USE OF AN OG/ON BASIS

Let $B = \{v_1, \dots, v_n\}$ be an OG set which is also a basis for V. Then for any $x \in V$,

$$x = \frac{\langle x, v_1 \rangle}{\|v_1\|^2} v_1 + \frac{\langle x, v_2 \rangle}{\|v_2\|^2} v_2 + \cdots + \frac{\langle x, v_n \rangle}{\|v_n\|^2} v_n$$

(So, the coefficients in the unique expression of x in terms of the basis vectors is easy to get, namely, they are $\frac{\langle x, v_i \rangle}{\|v_i\|^2}$.

Proof: B is a basis for V and $x \in V$.

$\Rightarrow \exists !$ ("there exists unique") $c_1, \dots, c_n \in \mathbb{R} \ni$

$$x = c_1 v_1 + c_2 v_2 + \dots + c_n v_n \quad \text{---} \quad (*)$$

To find c_1:

Take inner product with v_1 on both sides of $(*)$:

$$\langle x, v_1 \rangle = \langle c_1 v_1, v_1 \rangle + \langle c_2 v_2, v_2 \rangle + \dots + \langle c_n v_n, v_1 \rangle$$

(by additivity)

$$= c_1 \langle v_1, v_1 \rangle + c_2 \langle v_2, v_1 \rangle + \dots + c_n \langle v_n, v_1 \rangle$$

(by homogeneity)

$$= c_1 \|v_1\|^2 + \underbrace{0 + \dots + 0}$$

$$= c_1 \|v_1\|^2$$ (since $\langle v_i, v_j \rangle = 0 \ \forall \ i \neq j$ since S is an OG set)

$$\Rightarrow c_1 = \frac{\langle x, v_1 \rangle}{\|v_1\|^2}.$$

In general, $c_j = \frac{\langle x, v_j \rangle}{\|v_j\|^2}$

<u>Corollary</u>: If B is an ON basis, then $(*)$ becomes

$$x = \langle x, v_1 \rangle v_1 + \langle x, v_2 \rangle v_2 + \dots + \langle x_n, v_n \rangle v_n$$

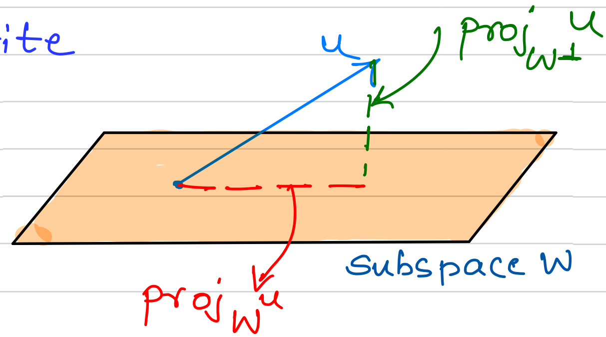
Eg. In $\mathbb{R}^3$, $\{e_1, e_2, e_3\}$ is an ON basis.

Q: Given a basis for V, can we make it into an OG basis? How?

For this, we first need the concept of projection

$\exists$ a unique way to write u as

$u = \text{proj}_W u + \text{proj}_{W^\perp} u$



$\text{proj}_W u$: orthogonal projection of u on W.

$\text{proj}_{W^\perp} v$: orthogonal projection of u perpendicular to W, or on $W^\perp$ (W perp)

Formula: If W has an OG basis $\{v_1, v_2, \dots, v_r\}$, then $\text{proj}_W u = \dfrac{\langle u, v_1 \rangle}{\|v_1\|^2} v_1 + \dots + \dfrac{\langle u, v_r \rangle}{\|v_r\|^2} v_r$

(So the formula we know for projections using dot product generalizes)

Sect. 6.3 – Gram - Schmidt process

Let $B = \{u_1, u_2, u_3\}$ be a basis for V.

AIM : Find a new basis $B' = \{v_1, v_2, v_3\}$ which is OG.

Solution :

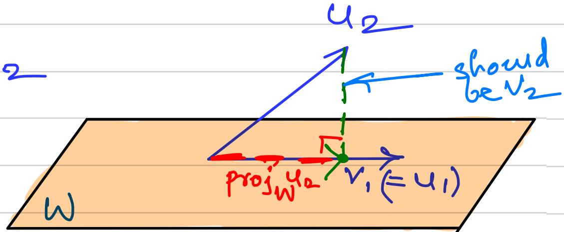
Step 1 : Take $v_1 = u_1$

Step 2 : To get v_2 :

(we can't take u_2, as u_2 may not be OG to v_1)

$\Rightarrow$

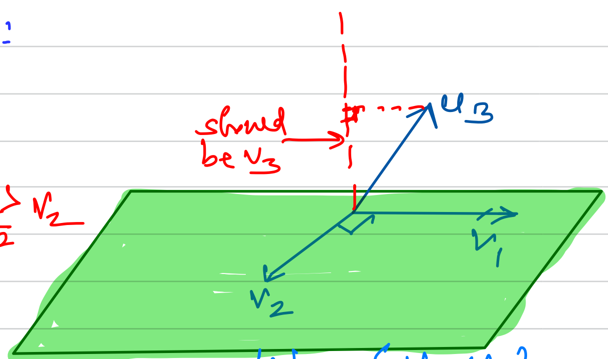
$$v_2 = u_2 - \hat{proj}_W u_2$$
$$= u_2 - \frac{\langle u_2, v_1 \rangle}{\|v_1\|^2} v_1$$



W has OG basis $\{v_1\}$

Step 3 : To get v_3 :

$$v_3 = u_3 - \hat{proj}_W u_3$$
$$= u_3 - \frac{\langle u_3, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle u_3, v_2 \rangle}{\|v_2\|^2} v_2$$



$W = \{v_1, v_2\}$

$\Rightarrow$ we get $\{v_1, v_2, v_3\} \longrightarrow$ OG set which will also be a basis for V

The Gram–Schmidt Process

To convert a basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ into an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$, perform the following computations:

Step 1. $\mathbf{v}_1 = \mathbf{u}_1$

Step 2. $\mathbf{v}_2 = \mathbf{u}_2 - \dfrac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$

Step 3. $\mathbf{v}_3 = \mathbf{u}_3 - \dfrac{\langle \mathbf{u}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \dfrac{\langle \mathbf{u}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$

Step 4. $\mathbf{v}_4 = \mathbf{u}_4 - \dfrac{\langle \mathbf{u}_4, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \dfrac{\langle \mathbf{u}_4, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 - \dfrac{\langle \mathbf{u}_4, \mathbf{v}_3 \rangle}{\|\mathbf{v}_3\|^2}\mathbf{v}_3$

$$\vdots$$

(continue for r steps)

Optional Step. To convert the orthogonal basis into an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_r\}$, normalize the orthogonal basis vectors.

i.e. $q_1 = \dfrac{v_1}{\|v_1\|}, \ldots, q_r = \dfrac{v_r}{\|v_r\|}$ where $v_1, \ldots, v_r$ are as obtained in Step 1 – r.

$\|q_1\| = \left\| \dfrac{v_1}{\|v_1\|} \right\| = \dfrac{1}{\|v_1\|} \|v_1\| = 1$

Two concepts:

1. Orthogonal complement:

Let $W \in V$ be a subspace. The orthogonal complement $W^\perp$ of W is

$$W^\perp = \{ x \in V : \langle x, w \rangle = 0 \ \forall \, w \in W \}$$

e.g. in $\mathbb{R}^3$ with $\langle x, y \rangle = x \cdot y$,
let $W = xz$-plane
$= \{ (x_1, 0, x_3) : x_1, x_3 \in \mathbb{R} \}$.

Then $W^\perp$?
want $y = (y_1, y_2, y_3)$

$y \cdot x = 0 \ \forall \, x \in W$
Take $x = (1, 0, 0) \implies y_1 = 0$
$x = (0, 0, 1) : \implies y_3 = 0$

$\implies W^\perp = \{ (0, y_2, 0) : y_2 \in \mathbb{R} \} = y$-axis.