1. <u>Orthogonal</u> <u>Vectors</u> :-

Two vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$ are orthogonal to each other if $\vec{u}.\vec{v} = 0$

ex $\vec{u} = (1,2)$ and $\vec{v} = (6,-3)$ are orthogonal in $\mathbb{R}^2$, as

$$\vec{u}.\vec{v} = (1,2).(6,-3) = 6-6 = 0$$

A set of vectors $\{\vec{v_1}, \vec{v_2}, \dots, \vec{v_n}\}$ are mutually orthogonal if every pair of vectors is orthogonal.

ie $\vec{v_i} \cdot \vec{v_j} = 0$, for all $i \neq j$.

The set of vectors $(1,0,-1)$, $(1,\sqrt{2},1)$, $(1,-\sqrt{2},1)$ are mutually orthogonal, since

$$(1,0,-1).(1,\sqrt{2},1) = 1+0-1 = 0$$
$$(1,0,-1),(1,-\sqrt{2},1) = 1+0-1 = 0$$
$$(1,\sqrt{2},1).(1,-\sqrt{2},1) = 1-2+1 = 0$$

3. <u>Orthogonal</u> <u>Subspaces</u>

Subspace S is orthogonal to subspace T means: every vector is S is orthogonal to every vector in T.

ex: In a plane, the space containing only the zero vector and any line through the origin are orthogonal subspaces.

A line through the origin and the whole plane are never orthogonal subspaces.

Two lines through the origin are orthogonal subspaces if they meet at right angles.

The row space of a matrix is orthogonal to
the nullspace, because $Ax = 0$ means the dot product
of x with each row of A is 0.
But then the product of x with any combination
of rows of A must be 0.

The columnspace is orthogonal to the left
nullspace of A because the rowspace of A^T is
perpendicular to the nullspace of A^T, as $A^T y = 0$.

$A_{m\times n}$
$N(A) \to \mathbb{R}^n$
$C(A) \to \mathbb{R}^m$
$R(A) \to \mathbb{R}^n$
$N(A^T) \to \mathbb{R}^m$
$A^T_{n\times m}$

ex: $A = \begin{bmatrix} 1 & 2 & 5 \\ 2 & 4 & 10 \end{bmatrix}_{2\times 3}$

$R_2 - 2R_1 \quad \sim \begin{bmatrix} 1 & 2 & 5 \\ 0 & 0 & 0 \end{bmatrix}$

Row space has dimension 1 with basis $\{(1, 2, 5)\}$

$Ax = 0 \Rightarrow \quad x_1 + 2x_2 + 5x_3 = 0 \Rightarrow x_1 = -2x_2 - 5x_3$

$$\therefore \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2x_2 - 5x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix}$$

Nullspace has dimension 2 with basis $\{(-2,1,0), (-5,0,1)\}$
which is orthogonal to the rowspace $\left\{ \begin{matrix} (-2,1,0)\cdot(1,2,5) = 0 \\ (-5,0,1)\cdot(1,2,5) = 0 \end{matrix} \right\}$

Not only is the nullspace orthogonal to the
rowspace, their dimensions add up to the dimension
of the whole space. The nullspace and the rowspace
are orthogonal complements in $\mathbb{R}^n$.
Similarly the column space and the left nullspace
are orthogonal complements in $\mathbb{R}^m$

Orthogonal complement:-

Let V be a subspace of $\mathbb{R}^n$.

The set $V^{\perp} = \{\vec{w} \mid \vec{w}\cdot\vec{v} = 0 \text{ for all } \vec{v} \in V\}$

is called the orthogonal complement of V.

Note:-

* A vector $\vec{w}$ is in $V^{\perp}$ iff $\vec{w}$ is orthogonal to every vector in a set that spans V.
* $V^{\perp}$ is a subspace of $\mathbb{R}^n$.

2. Orthogonal Sets:-

A set of vectors $\{u_1, \dots u_p\}$ in $\mathbb{R}^n$ is said to be an orthogonal set if each pair of distinct vectors from the set is orthogonal, that is, if $u_i \cdot u_j = 0$ whenever $i \neq j$.

ex $\{u_1, u_2, u_3\}$ such that $u_1 = (3,1,1)$, $u_2 = (-1,2,1)$,

$u_3 = (-\frac{1}{2}, -2, \frac{7}{2})$.

$u_1 \cdot u_2 = (3,1,1)\cdot(-1,2,1) = -3+2+1 = 0$

$u_1 \cdot u_3 = (3,1,1)\cdot(-\frac{1}{2},-2,\frac{7}{2}) = -\frac{3}{2} - 2 + \frac{7}{2} = 0$

$u_2 \cdot u_3 = (-1,2,1)\cdot(-\frac{1}{2},-2,\frac{7}{2}) = \frac{1}{2} - 4 + \frac{7}{2} = 0$

Each pair of distinct vectors is orthogonal, and so $\{u_1, u_2, u_3\}$ is an orthogonal set.

If $S = \{u_1, \dots, u_p\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is linearly independent and hence is a basis for the subspace spanned by S.

Proof: If $c_1u_1 + c_2u_2 + \dots + c_pu_p = 0$, for scalars $c_1, \dots, c_p$,

then $(c_1u_1 + c_2u_2 + \dots + c_pu_p) \cdot u_1 = 0 \cdot u_1$

$\Rightarrow (c_1u_1)\cdot u_1 + (c_2u_2)\cdot u_1 + \dots + (c_pu_p)\cdot u_1 = 0 \cdot u_1$

$\Rightarrow c_1(u_1 \cdot u_1) + c_2(u_2 \cdot u_1) + \dots + c_p(u_p \cdot u_1) = 0 \cdot u_1$

$\Rightarrow c_1(u_1 \cdot u_1) = 0$ $\left[\because u_2 \cdot u_1 = \dots = u_p \cdot u_1 = 0 \text{ as } \{u_1, \dots u_p\} \text{ is an orthogonal set}\right]$.

$\Rightarrow c_1 = 0$

Similarly $c_2 = \dots = c_p = 0$

$\therefore$ S is linearly independent.

4. Orthogonal basis:-

An orthogonal basis for a subspace W of $\mathbb{R}^n$ is a basis for W that is also an orthogonal set.

ex: $S = \{u_1, u_2, u_3\}$, $u_1 = (3,1,1)$, $u_2 = (-1,2,1)$, $u_3 = (-\frac{1}{2}, -2, \frac{7}{2})$ is an orthogonal basis for $\mathbb{R}^3$ as (i) S is an orthogonal set and (ii) S forms a basis of $\mathbb{R}^3$.

$$\begin{vmatrix} 3 & 1 & 1 \\ -1 & 2 & 1 \\ -\frac{1}{2} & -2 & \frac{7}{2} \end{vmatrix} = 3(7+2) - 1\left(-\frac{7}{2} + \frac{1}{2}\right) + 1(2+1) = 27 + 3 + 3 = 33 \neq 0$$

3. Orthonormal Sets

A set $\{u_1, \dots, u_p\}$ is an orthonormal set if it is an orthogonal set of unit vectors.
If W is the subspace spanned by such a set, then $\{u_1, \dots, u_p\}$ is an orthonormal basis for W, since the set is automatically linearly independent.

ex $\{e_1, \dots, e_n\}$, the standard basis for $\mathbb{R}^n$, is an orthonormal set.
Any nonempty subset of $\{e_1, \dots, e_n\}$ is orthonormal, too.

example:
5. Show that $\{v_1, v_2, v_3\}$ is an orthonormal basis of $\mathbb{R}^3$, where $v_1 = \left(\frac{3}{\sqrt{11}}, \frac{1}{\sqrt{11}}, \frac{1}{\sqrt{11}}\right)$, $v_2 = \left(-\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$, $v_3 = \left(-\frac{1}{\sqrt{66}}, -\frac{4}{\sqrt{66}}, \frac{7}{\sqrt{66}}\right)$

$$v_1 \cdot v_2 = -\frac{3}{\sqrt{66}} + \frac{2}{\sqrt{66}} + \frac{1}{\sqrt{66}} = 0$$

$$v_1 \cdot v_3 = -\frac{3}{\sqrt{726}} - \frac{4}{\sqrt{726}} + \frac{7}{\sqrt{726}} = 0$$

$$v_2 \cdot v_3 = \frac{1}{\sqrt{396}} - \frac{8}{\sqrt{396}} + \frac{7}{\sqrt{396}} = 0$$

Thus $\{v_1, v_2, v_3\}$ is an orthogonal set.

$$v_1 \cdot v_2 = \frac{9}{\cancel{\sqrt{11}}} + \frac{1}{11} + \frac{1}{11} = 1$$

$$v_2 \cdot v_2 = \frac{1}{6} + \frac{4}{6} + \frac{1}{6} = 1$$

$$v_3 \cdot v_3 = \frac{1}{66} + \frac{16}{66} + \frac{49}{66} = 1$$

which shows that v_1, v_2 and v_3 are unit vectors.
Thus $\{v_1, v_2, v_3\}$ is an orthonormal set.
Since the set is linearly independent, its three vectors form a basis for $\mathbb{R}^3$.

ex: Show that $\{u_1, u_2\}$, where $u_1 = \left(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$

$u_2 = \left(\frac{2}{\sqrt{5}}\ \frac{1}{\sqrt{5}}\right)$ is an orthonormal basis for $\mathbb{R}^2$.

6. Orthogonal matrix

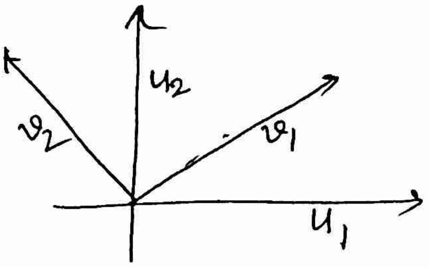
A ~~square~~ matrix A with real entries and satisfying the condition $A^{-1} = A^T$ is called an orthogonal matrix.

The vectors $u_1 = (1,0)$ and $u_2 = (0,1)$ form an orthonormal basis $B = \{u_1, u_2\}$.

Rotating the vectors u_1 and u_2 anticlockwise by an ange θ, we obtain $R_\theta = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$

$v_1 = (\cos\theta, \sin\theta)$ and $v_2 = (-\sin\theta, \cos\theta)$.

Then $C = \{v_1, v_2\}$ is also an orthonormal basis



The transition matrix from the basis C to the basis B is given by

$$\left[\begin{array}{cc:cc} 1 & 0 & \cos\theta & -\sin\theta \\ 0 & 1 & \sin\theta & \cos\theta \end{array}\right]$$

$$P_{B\leftarrow C} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

~~Clearly~~ $P^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$ $\quad P^T = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$

Clearly $P^{-1} = P^T$.

$\therefore$ P is an orthogonal matrix.

* Suppose that $B = \{u_1, \ldots, u_n\}$ and $C = \{v_1, \ldots, v_n\}$ are two orthonormal bases of a vectorspace V. Then the transition matrix P from the basis C to the basis B is an orthogonal matrix.

example:

The matrix $A = \begin{bmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix}$ is orthogonal,

Since $A^T A = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

The row vector of A, namely $(\frac{1}{3}, -\frac{2}{3}, \frac{2}{3})$, $(\frac{2}{3}, -\frac{1}{3}, -\frac{2}{3})$ and $(\frac{2}{3}, \frac{2}{3}, \frac{1}{3})$ are orthonormal.

So are the column vectors of A.

* Suppose that A is an $n \times n$ matrix with real entries. Then (a) A is orthogonal iff the row vectors of A form an orthonormal basis of $\mathbb{R}^n$.

(b) A is orthogonal iff the column vectors of A form an orthonormal basis of $\mathbb{R}^n$.

ex:- Show that the matrix $U = \begin{bmatrix} \frac{3}{\sqrt{11}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{66}} \\ \frac{1}{\sqrt{11}} & \frac{2}{\sqrt{6}} & -\frac{4}{\sqrt{66}} \\ \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{6}} & \frac{7}{\sqrt{66}} \end{bmatrix}$ is an orthogonal matrix.

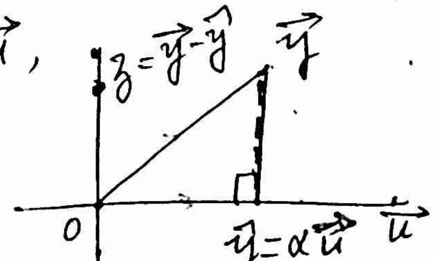
$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

7. Orthogonal Projections:

Given a nonzero vector $\vec{u}$ in $\mathbb{R}^n$, consider the problem of decomposing a vector $\vec{y}$ in $\mathbb{R}^n$ into the sum of two vectors, one a multiple of $\vec{u}$ and the other orthogonal to $\vec{u}$. We wish to write

$$\vec{y} = \overset{\alpha\vec{u}}{\hat{y}} + \vec{z} \quad (1)$$

where $\hat{y} = \alpha\vec{u}$ for some scalar α and $\vec{z}$ is some vector orthogonal to $\vec{u}$.



Finding α to make $y - \hat{y}$ orthogonal to $\vec{u}$.

Given any scalar α, let $\vec{z} = \vec{y} - \alpha\vec{u}$, so that (1) is satisfied.

Then $\overset{\vec{z}}{\vec{y} - \hat{y}}$ is orthogonal to $\vec{u}$ iff

$$0 = (\vec{y} - \alpha\vec{u}) \cdot \vec{u} = \vec{y} \cdot \vec{u} - (\alpha\vec{u}) \cdot \vec{u}$$
$$= \vec{y} \cdot \vec{u} - \alpha(\vec{u} \cdot \vec{u})$$

That is, (1) is satisfied with $\vec{z}$ orthogonal to $\vec{u}$ iff $\alpha = \dfrac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}}$ and $\overset{\alpha\vec{u}}{\hat{y}} = \dfrac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}}\vec{u}$.

The vector $\overset{\text{denoted as } \hat{y}}{\alpha\vec{u} = \hat{y}}$ is called the orthogonal projection of $\vec{y}$ onto $\vec{u}$, and the vector $\vec{z}$ is called the component of $\vec{y}$ orthogonal to $\vec{u}$.

ex:- Let $\vec{y} = (7,6)$ and $\vec{u} = (4,2)$.

Find the orthogonal projection of $\vec{y}$ onto $\vec{u}$.
Then write $\vec{y}$ as the sum of two orthogonal vectors, one in Span$\{\vec{u}\}$ and one orthogonal to $\vec{u}$.

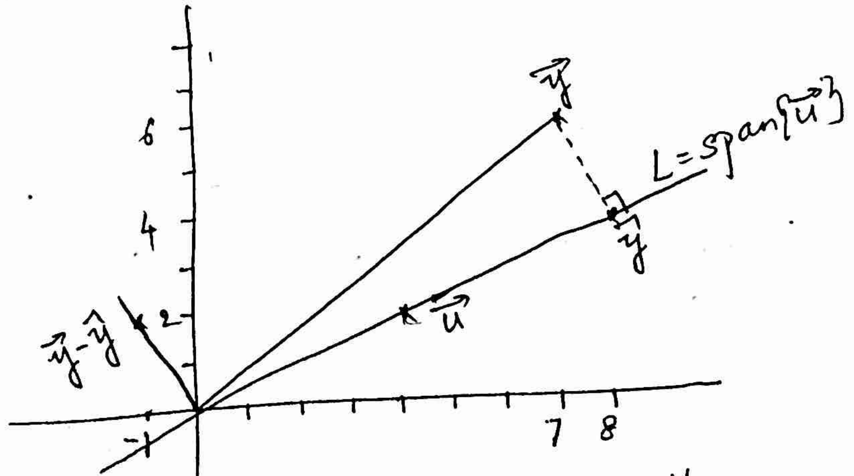
Solⁿ:
$$\vec{y}\cdot\vec{u} = (7,6)\cdot(4,2) = 28+12 = 40$$
$$\vec{u}\cdot\vec{u} = (4,2)\cdot(4,2) = 16+4 = 20.$$

The orthogonal projection of $\vec{y}$ onto $\vec{u}$ is
$$\hat{y} = \frac{\vec{y}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\vec{u} = \frac{40}{20}\vec{u} = 2\vec{u} = 2(4,2) = (8,4).$$

The component of $\vec{y}$ orthogonal to $\vec{u}$ is
$$\vec{y} - \hat{y} = (7,6) - (8,4) = (-1,2)$$

The component of $\vec{y}$ in span$\{\vec{u}\}$ is
$$\alpha\vec{u} = 2(4,2) = (8,4).$$

$$\therefore \quad \vec{y} = \alpha\vec{u} + (\vec{y} - \hat{y})$$
$$= (8,4) + (-1,2)$$

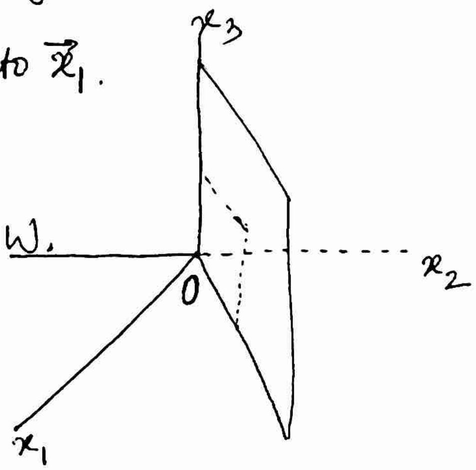


ex: Let $\vec{y} = (2,3)$ and $\vec{u} = (4,-7)$. Write $\vec{y}$ as the sum of two orthogonal vectors, one in span$\{\vec{u}\}$ and a vector orthogonal to $\vec{u}$.

Gram-Schmidt Orthogonalization.

The Gram-Schmidt process is a single algorithm for producing an orthogonal or orthonormal basis for any nonzero subspace of $\mathbb{R}^n$.

ex! Let $W = \text{Span}\{x_1, x_2\}$ where $x_1 = (3,6,0)$ and $x_2 = (1,2,2)$. Construct an orthogonal basis $\{v_1, v_2\}$ for W.



Let $\vec{p}$ be the projection of $\vec{x}_2$ onto $\vec{x}_1$.
The component of $\vec{x}_2$ orthogonal to $\vec{x}_1$ is $\vec{x}_2 - \vec{p}$, which is in W.
Let $\vec{v}_1 = \vec{x}_1$ and

$$\vec{v}_2 = \vec{x}_2 - \vec{p}$$

$$= \vec{x}_2 - \frac{\vec{x}_2 \cdot \vec{x}_1}{\vec{x}_1 \cdot \vec{x}_1}\vec{x}_1$$

$$= (1,2,2) - \frac{(1,2,2)\cdot(3,6,0)}{(3,6,0)\cdot(3,6,0)}(3,6,0)$$

$$\vec{v}_2 = (0,0,2).$$

Then $\{\vec{v}_1, \vec{v}_2\}$ is an orthogonal set of nonzero vectors in W. Since $\dim W = 2$, the set $\{\vec{v}_1, \vec{v}_2\}$ is a basis in W.

ex: Let $W = \text{Span}\{v_1, v_2\}$ where $v_1 = (1,1)$ & $v_2 = (2,-1)$.

Construct an orthogonal basis $\{u_1, u_2\}$ for W.

$u_1 = v_1$

Sol^n: Set $u_1 = (1,1)$

and
$$u_2 = v_2 - \frac{v_2 \cdot u_1}{u_1 \cdot u_1} u_1$$
$$= (2,-1) - \frac{(2,-1)\cdot(1,1)}{(1,1)\cdot(1,1)}(1,1)$$
$$= \left(\frac{3}{2}, -\frac{3}{2}\right)$$

ex: Let $W = \text{Span}\{v_1, v_2, v_3\}$, where $v_1 = (0,1,2)$, $v_2 = (1,1,2)$, $v_3 = (1,0,1)$. Construct an orthogonal basis $\{u_1, u_2, u_3\}$ for W.

$u_1 = v_1$

Sol^n: Set $u_1 = (0,1,2)$

and
$$u_2 = v_2 - \frac{v_2 \cdot u_1}{u_1 \cdot u_1} u_1$$
$$= (1,1,2) - \frac{(1,1,2)\cdot(0,1,2)}{(0,1,2)\cdot(0,1,2)}(0,1,2)$$
$$u_2 = (1,0,0)$$

and
$$u_3 = v_3 - \frac{v_3 \cdot u_1}{u_1 \cdot u_1} u_1 \cancel{+} - \frac{v_3 \cdot u_2}{u_2 \cdot u_2} u_2$$
$$= (1,0,1) - \frac{(1,0,1)\cdot(0,1,2)}{(0,1,2)\cdot(0,1,2)}(0,1,2) - \frac{(1,0,1)\cdot(1,0,0)}{(1,0,0)\cdot(1,0,0)}(1,0,0)$$
$$= (1,0,1) - \frac{2}{5}(0,1,2) - (1,0,0)$$
$$u_3 = \left(0, -\frac{2}{5}, \frac{1}{5}\right)$$

The Gram-Schmidt Process [gram-shmit]

Given a basis $\{x_1, \dots, x_p\}$ for a subspace W of $\mathbb{R}^n$, define

$$v_1 = x_1$$

$$v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$$

$$v_3 = x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2$$

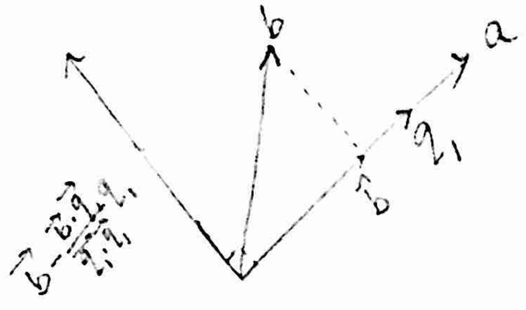
$$\vdots$$

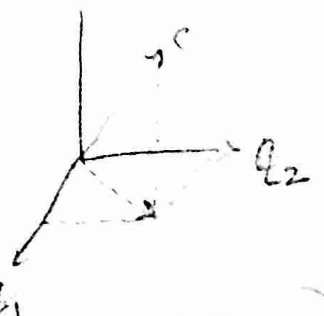
$$v_p = v_p - \frac{x_p \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_p \cdot v_2}{v_2 \cdot v_2} v_2 - \cdots - \frac{x_p \cdot v_{p-1}}{v_{p-1} \cdot v_{p-1}} v_{p-1}$$

Then $\{v_1, \dots, v_p\}$ is an orthogonal basis for W.

In addition, $\mathrm{Span}\{v_1, \dots v_k\} = \mathrm{Span}\{x_1, \dots x_k\}$ for $1 \le k \le p$.

The construction, which converts a skewed set of axes into a perpendicular set, is known as Gram-Schmidt Orthogonalization.





$$\vec{c} - \left(\frac{\vec{c} \cdot q_1}{q_1 \cdot q_1} q_1 + \frac{\vec{c} \cdot q_2}{q_2 \cdot q_2} q_2 \right)$$

example:

Find an orthogonal basis for the column space of the matrix $\begin{bmatrix} 3 & -5 & 1 \\ 1 & 1 & 1 \\ -1 & 5 & -2 \\ 3 & -7 & 8 \end{bmatrix}$

Sol: The columns of A are the vectors $\{x_1, x_2, x_3\}$

$x_1 = (3,1,-1,3)$, $x_2 = (-5,1,5,-7)$

$x_3 = (1,1,-2,8)$

Let $v_1 = (3,1,-1,3)$

$$v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$$

$$= (-5,1,5,-7) - \frac{(-5,1,5,-7)\cdot(3,1,-1,3)}{(3,1,-1,3)\cdot(3,1,-1,3)}(3,1,-1,3)$$

$$= (-5,1,5,-7) - \frac{(-40)}{(20)}(3,1,-1,3)$$

$-5+6 \mid 1+2$

$5-2 \mid -7+6$

$v_2 = (1,3,3,-1)$

$$v_3 = x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1}(v_1) \mp \frac{x_3 \cdot v_2}{v_2 \cdot v_2}(v_2)$$

$$= (1,1,-2,8) - \frac{(1,1,-2,8)\cdot(3,1,-1,3)}{(3,1,-1,3)\cdot(3,1,-1,3)}(3,1,-1,3) - \frac{(1,1,-2,8)\cdot(1,3,3,-1)}{(1,3,3,-1)\cdot(1,3,3,-1)}(1,3,3,-1)$$

$$= (1,1,-2,8) - \frac{30}{20}(3,1,-1,3) - \frac{(-10)}{20}(1,3,3,-1)$$

$1 - \frac{9}{2} + \frac{1}{2} = -\frac{6}{2}$

$1 - \frac{3}{2} + \frac{3}{2} = 1$

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

$8 - \frac{9}{2} - \frac{1}{2} = \frac{6}{2}$

$v_3 = (-3,1,1,3)$

$\{(3,1,-1,3), (1,3,3,-1), (-3,1,1,3)\}$ is an orthogonal basis for the column space of the given matrix.

example

Find an orthogonal basis for the column space of the matrix

$$A = \begin{bmatrix} -1 & 6 & 6 \\ 3 & -8 & 3 \\ 1 & -2 & 6 \\ 1 & -4 & -3 \end{bmatrix}$$

The columns of A are $\{x_1, x_2, x_3\}$,

where $x_1 = (-1,3,1,1)$, $x_2 = (6,-8,-2,-4)$, $x_3 = (6,3,6,-3)$

let $v_1 = (-1,3,1,1)$

$$v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$$

$$= (6,-8,-2,-4) - \frac{(6,-8,-2,-4)\cdot(-1,3,1,1)}{(-1,3,1,1)\cdot(-1,3,1,1)}(-1,3,1,1)$$

$$= (6,-8,-2,-4) - \frac{(\cancel{-36})^{3}}{\cancel{12}}(-1,3,1,1)$$

$-6-24-2-4$

$1+9+1+1$

$6-3,\ -8+9$

$-2+3,\ -4+3$

$$v_2 = (3,1,1,-1)$$

$$v_3 = x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2,$$

$$= (6,3,6,-3) - \frac{(6,3,6,-3)\cdot(-1,3,1,1)}{(-1,3,1,1)\cdot(-1,3,1,1)}(-1,3,1,1) - \frac{(6,3,6,-3)\cdot(3,1,1,-1)}{(3,1,1,-1)\,(3,1,1,-1)}(3,1,1,-1)$$

$$= (6,3,6,-3) - \frac{\cancel{6}}{\cancel{12}_2}(-1,3,1,1) - \frac{\cancel{30}^{5}}{\cancel{12}_2}(3,1,1,-1)$$

$$= (6,3,6,-3) - \left(-\tfrac{1}{2}, \tfrac{3}{2}, \tfrac{1}{2}, \tfrac{1}{2}\right) - \left(\tfrac{15}{2}, \tfrac{5}{2}, \tfrac{5}{2}, -\tfrac{5}{2}\right)$$

$$v_3 = \left(\cancel{-\tfrac{5}{2}}, \cancel{-\tfrac{3}{2}}, \cancel{\tfrac{5}{2}}, \cancel{\tfrac{1}{2}}\right)\ (-1,-1,3,-1)$$

$-6+9+6-3$

$18+3+6+3$

$9+1+1-1$

$6+\frac{1}{2}-\frac{15}{2} = \cancel{\frac{12+1-15}{2}}$

$3-\frac{3}{2}-\frac{5}{2}$

$6-\frac{1}{2}-\frac{5}{2} = \frac{12-1-6}{2}$

$-3-\frac{1}{2}+\frac{5}{2} = \cancel{-6-1+6}$

$\{(-1,3,1,1), (3,1,1,-1), (-1,-1,3,-1)\}$ is an orthogonal basis for the column space of the given matrix.

Least - Squares Problems

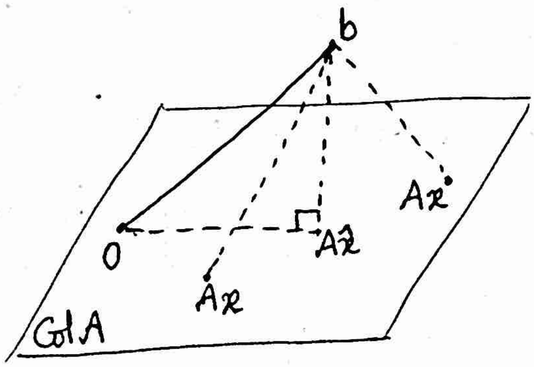
Suppose ~~that~~ the system $Ax=b$ is inconsistent, i.e, the solution doesnot exist. The best one can do is to find an x that makes Ax as close as possible to b.

Think of Ax as an approximation to b. The smaller the distance between b and Ax, given by $|b-Ax|$, the better the approximation. The general least-squares problem is to find an x that makes $|b-Ax|$ as small as possible.

Definition:-

If A is $m\times n$ and b is in $\mathbb{R}^m$, a least-squares solution of $Ax=b$ is an $\hat{x}$ in $\mathbb{R}^n$ such that $|b-A\hat{x}|\leq|b-Ax|$ for all x in $\mathbb{R}^n$.

No matter what x we select, the vector Ax will necessarily be in the column space of A. So we seek an x that makes Ax the closest point in $\mathrm{Col}A$ to b.



$A_{m\times n}$

$x_{n\times 1}$

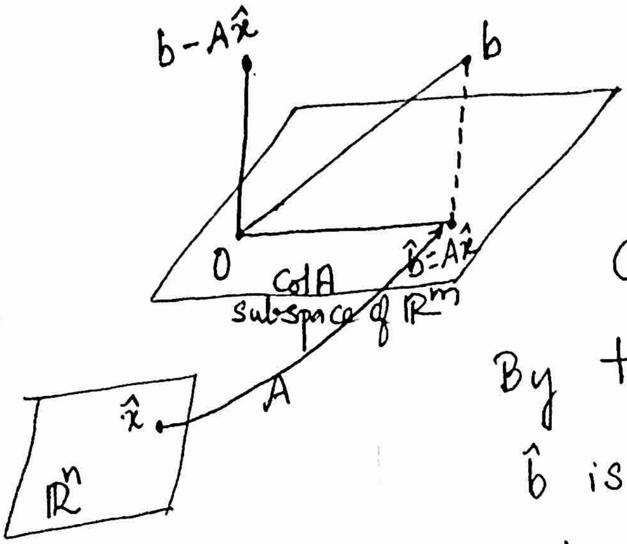
$b_{m\times 1}$

$(A\hat{x})_{m\times 1}$

The vector b is closer to $A\hat{x}$ than to Ax.

$\mathrm{Col}\,A \in \mathbb{R}^m$

Solution of the General Least-Squares Problem



Let A be an $m \times n$ matrix and b is in $\mathbb{R}^m$

Col A is a subspace of $\mathbb{R}^m$

By the Best Approximation theorem, $\hat{b}$ is the orthogonal projection of b onto Col A.

Because $\hat{b}$ is in the column space of A, the equation $Ax = \hat{b}$ is consistent, and there is an $\hat{x}$ in $\mathbb{R}^n$ such that $A\hat{x} = \hat{b}$. — ①

Since $\hat{b}$ is ~~is~~ the closest point in Col A to b, a vector $\hat{x}$ is a least-squares solution of $Ax = b$ iff $\hat{x}$ satisfies ①.

Suppose $\hat{x}$ satisfies $A\hat{x} = \hat{b}$.

By the Orthogonal Decomposition Theorem, the projection $\hat{b}$ which lies in Col A is orthogonal to $b - \hat{b}$, i.e, $b - \hat{b}$ is orthogonal to each column of A.

i.e, $b - A\hat{x}$ is orthogonal to each column of A.

If a_j is any column of A, then $a_j \cdot (b - A\hat{x}) = 0$, and $a_j^T (b - A\hat{x}) = 0$. Since each a_j^T is a row of A^T,

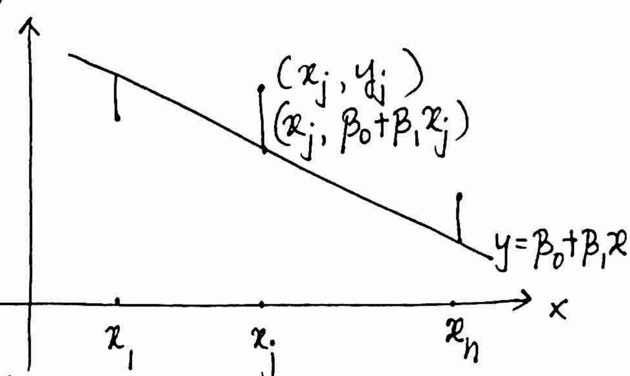
$A^T(b - A\hat{x}) = 0$ — ② $\Rightarrow$ $A^T b - A^T A\hat{x} = 0 \Rightarrow A^T A\hat{x} = A^T b$

These calculations show that each least-squares sol^n of $Ax = b$ satisfies the equation $A^T A x = A^T b$ — ③, whose sol^n is denoted by $\hat{x}$.

The matrix equ^n ③ represents a system of equ^s called normal eq^s for $Ax = b$

Applications to linear models

Suppose we want to fit a linear equation of the form $y = \beta_0 + \beta_1 x$ to the data points $(x_1, y_1) \dots, (x_n, y_n)$, that, when graphed, seem to lie close to a line. We want to determine the parameters β_0 and β_1 that make the line as close to the points as possible.



Suppose β_0 and β_1 are fixed and consider the line $y=\beta_0+\beta_1 x$. Corresponding to each data point (x_j, y_j) there is a point $(x_j, \beta_0+\beta_1 x_j)$ on the line with the same x-coordinates.

We call y_j the observed value of y and $\beta_0+\beta_1 x_j$ the predicted y-value. The difference between an observed y-value and a predicted y-value is called a residual. The best fit line to ~~thes~~ given data is such that, the square of these residuals is minimum. And that is the least-square line $y = \beta_0 + \beta_1 x$. This line is also called a line of regression of y on x, because any errors in the data are assumed to be only in the y-coordinates. The coefficients β_0, β_1 of the line are called regression coefficients.

If the data points were on the line, the parameters β_0 and β_1 would satisfy the equations $\longrightarrow$

predicted y-value		observed y-value
$\beta_0 + \beta_1 x_1$	$=$	y_1
$\beta_0 + \beta_1 x_2$	$=$	y_2
		$\vdots$
$\beta_0 + \beta_1 x_n$	$=$	y_n

We can write the system as

$$X\beta = y, \quad \text{where} \quad X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

If the data points don't lie on a line, then there are no parameters β_0, β_1 for which the predicted y-values in $X\beta$ equal the observed y-values in y and $X\beta = y$ has no solution. This is a least-squares problem $Ax = b$, with different notation.

The square of the distance between the vectors $X\beta$ and y is precisely the sum of the squares of the residuals. The β that minimizes the sum also minimizes the distance between $X\beta$ and y.

Computing the least-squares solution of $X\beta = y$ is equivalent to finding the β that determines the least-squares line ($y = \beta_0 + \beta_1 x$).

Theorem

The set of least-squares solutions of $Ax = b$ coincides with the nonempty set of solutions of the normal equations $A^TAx = A^Tb$.

example:-

Find a least-squares solution of the inconsistent system $Ax = b$ for $A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}$, $b = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$

Sol^n:

$$A^TA = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$

$$A^Tb = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Then the equation $A^TAx = A^Tb$ becomes

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

$$\xrightarrow{17R_2 - R_1} \begin{bmatrix} 17 & 1 \\ 0 & 84 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 168 \end{bmatrix}$$

$$\Rightarrow \quad 17x_1 + x_2 = 19 \quad \Longrightarrow \quad x_2 = 2$$

$$84x_2 = 168 \qquad\qquad x_1 = 1$$

$\therefore$ $\hat{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is the least-squares solution.

example:-

Find a least-squares solution of $Ax=b$ for

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \qquad b = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix}$$

Sol^n:-

$$A^TA = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}$$

$$A^Tb = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 2 \\ 6 \end{bmatrix}$$

Then $A^TAx = Ab \Rightarrow$

$$\begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 2 \\ 6 \end{bmatrix}$$

The augmented matrix $[A^TA : Ab] \Rightarrow$

$$\left[\begin{array}{cccc:c} 6 & 2 & 2 & 2 & 4 \\ 2 & 2 & 0 & 0 & -4 \\ 2 & 0 & 2 & 0 & 2 \\ 2 & 0 & 0 & 2 & 6 \end{array}\right] \overset{3R_2-R_1;\ 3R_3-R_1;\ 3R_4-R_1}{\sim} \left[\begin{array}{cccc:c} 6 & 2 & 2 & 2 & 4 \\ 0 & 4 & -2 & -2 & -16 \\ 0 & -2 & 4 & -2 & 2 \\ 0 & -2 & -2 & 4 & 14 \end{array}\right] \overset{2R_3+R_2;\ 2R_4+R_2}{\sim} \left[\begin{array}{cccc:c} 6 & 2 & 2 & 2 & 4 \\ 0 & 4 & -2 & -2 & -16 \\ 0 & 0 & 6 & -6 & -12 \\ 0 & 0 & -6 & 6 & 12 \end{array}\right]$$

$$\overset{R_4+R_3}{\sim} \left[\begin{array}{cccc:c} 6 & 2 & 2 & 2 & 4 \\ 0 & 4 & -2 & -2 & -16 \\ 0 & 0 & 6 & -6 & -12 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow \left.\begin{aligned} 6x_1+2x_2+x_3+2x_4 &= 4 \\ 4x_2-2x_3-2x_4 &= -16 \\ 6x_3-6x_4 &= -12 \end{aligned}\right\} \Rightarrow \begin{aligned} x_3 &= -2+x_4 \\ x_2 &= -5+x_4 \\ x_1 &= 3-x_4 \end{aligned} \quad \therefore \hat{x} = \begin{bmatrix} 3-x_4 \\ -5+x_4 \\ -2+x_4 \\ x_4 \end{bmatrix}$$

~~let $x_4=0 \Rightarrow x_3=-2,\ x_2=-5,\ x_1=3$~~

$$\hat{x} = \begin{bmatrix} 3 \\ -5 \\ -2 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

example:

Find the equation $y = \beta_0 + \beta_1 x$ of the least-squares line that best fits the data points $(2,1)$, $(5,2)$, $(7,3)$, $(8,3)$.

Sol^n:

Using the x-coordinates and y-coordinates of the data points, we can write

$$X = \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix}, \quad y = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}$$

For the least-squares solution $X\beta = y$, we obtain the normal equations by $X^T X \beta = X^T y$.

ie,
$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

$$X^T y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

The normal equations are

$$\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

Hence
$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}^{-1} \begin{bmatrix} 9 \\ 57 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 142 & -22 \\ -22 & 4 \end{bmatrix} \begin{bmatrix} 9 \\ 57 \end{bmatrix} = \begin{bmatrix} \frac{2}{7} \\ \frac{5}{14} \end{bmatrix}$$

Thus the least-squares line has the equation

$$y = \frac{2}{7} + \frac{5}{14} x$$

example:

Find the equation $y = \beta_0 + \beta_1 x$ of the least-squares line that best fits the data points $(-1, 0), (0, 1), (1, 2), (2, 4)$.

Sol: From the data points, we have,

$$X = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad y = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 4 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix}$$

0+1+2+4

0+0+2+8

$$X^T y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ 10 \end{bmatrix}$$

The normal equ^s are $X^T X \beta = X^T y$

24 − 4

42 − 20 = 22/20

−14 + 40 = 26/20

ie,
$$\begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 7 \\ 10 \end{bmatrix}$$

Hence
$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix}^{-1} \begin{bmatrix} 7 \\ 10 \end{bmatrix} = \frac{1}{20} \begin{bmatrix} 6 & -2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 7 \\ 10 \end{bmatrix} = \begin{bmatrix} \frac{11}{20} \\ \frac{13}{10} \end{bmatrix}$$

$\therefore \quad \underline{y = \frac{11}{10} + \frac{13}{10} x}$ is the required least-squares line.

example:

Find the line of best fit for the below data:

$b=2$ at $t=-1$, $b=0$ at $t=0$, $b=-3$ at $t=1$,

$b=-5$ at $t=2$.

Sol^n: Given $(-1,2)$, $(0,0)$, $(1,-3)$, $(2,-5)$.

Using the data points, we can write,

$$X = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad y = \begin{bmatrix} 2 \\ 0 \\ -3 \\ -5 \end{bmatrix}$$

For the least squares solution $X\beta = y$, we obtain the normal equations by $X^T X \beta = X^T y$

ie $$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix}$$

$$X^T y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ -3 \\ -5 \end{bmatrix} = \begin{bmatrix} -6 \\ -15 \end{bmatrix}$$

The normal equations are

$$\begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} -6 \\ -15 \end{bmatrix}$$

Hence $$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix}^{-1} \begin{bmatrix} -6 \\ -15 \end{bmatrix} = \frac{1}{20} \begin{bmatrix} 6 & -2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} -6 \\ -15 \end{bmatrix} = \begin{bmatrix} -\frac{6}{20} \\ -\frac{48}{20} \end{bmatrix} = \begin{bmatrix} -\frac{3}{10} \\ -\frac{24}{10} \end{bmatrix}$$

Thus the least-squares or the line of best fit is

$$y = -\frac{3}{10} - \frac{24}{10} x \quad \text{or} \quad \boxed{b = -0.3 - 2.4t}$$

A healthy child's systolic blood pressure p (in millimeters of mercury) and weight w (in pounds) are approximately related by the equation

$\beta_0 + \beta_1 \ln w = p$

Use the following experimental data to estimate the systolic blood pressure of a healthy child weighing 100 pounds

w	44	61	81	113	131
$\ln w$	3.78	4.11	4.39	4.73	4.88
p	91	98	103	110	112

$X_{5\times 2}$ $X^T_{2\times 5}$ $y_{5\times 1}$

$$X^T X = \begin{bmatrix} 5 & 21.89 \\ 21.89 & 96.6399 \end{bmatrix}$$

$$X^T y = \begin{bmatrix} 514 \\ 2265.8 \end{bmatrix}$$

$$X^T X \beta = X^T y \quad \begin{bmatrix} & \end{bmatrix}\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \ \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = [X^T X]^{-1} [X^T y]$$

$$= \begin{matrix} 18.5642 \\ 19.2407 \end{matrix}$$

Eigen values and Eigen vectors

If A is a square matrix of order n, we can find the matrix $A-\lambda I$, where I is the n^{th} order unit matrix. The determinant of this matrix equated to zero, i.e.,

$$|A-\lambda I| = \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} = 0$$

is called the characteristic equation of A.

On expanding the determinant, the characteristic equation takes the form

$$(-1)^n \lambda^n + k_1 \lambda^{n-1} + k_2 \lambda^{n-2} + \cdots + k_n = 0 ,$$

where $k^{'s}$ are expressible in terms of the elements a_{ij}. The roots of this equation are called the characteristic roots or latent roots or eigen-values of the matrix A.

$\frac{dv}{dt} = 4v - 5w$, $v = 8$ at $t = 0$

$\frac{dw}{dt} = 2v - 3w$, $w = 5$ at $t = 0$

$u(t) = \begin{bmatrix} v(t) \\ w(t) \end{bmatrix}$, $u(0) = \begin{bmatrix} 8 \\ 5 \end{bmatrix}$, $A = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix}$

$\frac{du}{dt} = Au$, $u = u(0)$ at $t = 0$

$v(t) = e^{\lambda t} x_1$

$w(t) = e^{\lambda t} x_2$

$u(t) = e^{\lambda t} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$\lambda e^{\lambda t} x_1 = 4 e^{\lambda t} x_1 - 5 e^{\lambda t} x_2$

$\lambda e^{\lambda t} x_2 = 2 e^{\lambda t} x_1 - 3 e^{\lambda t} x_2$

$4x_1 - 5x_2 = \lambda x_1$

$2x_1 - 3x_2 = \lambda x_2$

$Ax = \lambda x$

$(A - \lambda I)x = 0$

If $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$

then the linear transformation $Y = AX$ - (1) carries the column vector X into the column vector Y by means of the square matrix A.

In practice it is often required to find such vectors which transform into themselves or to a scalar multiple of themselves.

Let X be such a vector which transforms into λX by means of the transformation (1).

Then, $\lambda X = AX$ or $AX - \lambda IX = 0$ or $[A - \lambda I]X = 0$ - (2)

The matrix equation represents n homogeneous linear equations,

$$\left.\begin{aligned} (a_{11}-\lambda)x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + (a_{22}-\lambda)x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + (a_{nn}-\lambda)x_n &= 0 \end{aligned}\right\} \quad - (3)$$

which will have a non-trivial solution only if the coefficient matrix is singular.

i.e, if $|A - \lambda I| = 0$

This is called the characteristic equation of the transformation and is same as the characteristic equation of the matrix A.

It has n roots and corresponding to each root, the equation (2) (or equation (3)) will have a non-zero solution, $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, which is known as the eigen vector or latent vector.

Observation 1:-

Corresponding to n distinct eigen values, we get n independent eigen vectors. But when two or more eigen values are equal, it may or may not be possible to get linearly independent eigen vectors corresponding to the repeated roots.

Observation 2:-

If X_i is a solution for a eigen value λ_i then it follows from (2) that cX_i is also a solution, where c is an arbitrary constant. Thus the eigen vector corresponding to an eigen value is not unique, but may be any one of the vectors cX_i

Problems

1. Find the eigen values and eigen vectors of the matrix $\begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}$.

Solution:-

The characteristic equation is $|A-\lambda I|=0$.

ie $\begin{vmatrix} 5-\lambda & 4 \\ 1 & 2-\lambda \end{vmatrix} = 0$

$\Rightarrow (5-\lambda)(2-\lambda) - 4 = 0$

$\Rightarrow 10 - 5\lambda - 2\lambda + \lambda^2 - 4 = 0$

$\Rightarrow \lambda^2 - 7\lambda + 6 = 0$

$\Rightarrow (\lambda - 6)(\lambda - 1) = 0$

$\Rightarrow \underline{\lambda = 1, \ \lambda = 6}$ are the eigen values.

If $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ is an eigen vector corresponding to the eigen value λ, then

$[A - \lambda I] X = 0$

$\Rightarrow \begin{bmatrix} 5-\lambda & 4 \\ 1 & 2-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

For $\lambda = 1$, we have

$\begin{bmatrix} 5-1 & 4 \\ 1 & 2-1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\begin{bmatrix} 4 & 4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\Rightarrow \left.\begin{matrix} 4x_1 + 4x_2 = 0 \\ x_1 + x_2 = 0 \end{matrix}\right\}$.

or $x_1 = -x_2$

let $x_2 = k$, then $x_1 = -k$.

$\therefore X = \begin{bmatrix} -k \\ k \end{bmatrix}$ is the eigen vector corresponding to $\lambda = 1$.

For $\lambda = 6$, we have

$$\begin{bmatrix} 5-6 & 4 \\ 1 & 2-6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 4 \\ 1 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$\Rightarrow \quad -x_1 + 4x_2 = 0$

or $\quad x_1 = 4x_2$

Let $x_2 = k$, then $x_1 = 4k$

$\therefore \; X = \begin{bmatrix} 4k \\ k \end{bmatrix}$ is the eigen vector corresponding to $\lambda = 6$.

2. Find the eigen values and eigenvectors of the matrix $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$.

Solution:-

The characteristic equation is $|A - \lambda I| = 0$

i.e, $\begin{vmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{vmatrix} = 0$

$\Rightarrow \quad (1-\lambda)(4-\lambda) - 4 = 0$

$\Rightarrow \quad 4 - \lambda - 4\lambda + \lambda^2 - 4 = 0$

$\Rightarrow \quad \lambda^2 - 5\lambda = 0$

$\Rightarrow \quad \lambda(\lambda - 5) = 0$

$\Rightarrow \quad \underline{\lambda = 0, \;\; \lambda = 5}$ are the eigen values.

$\longrightarrow$

If $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ is an eigen vector corresponding to the eigen value λ, then $[A - \lambda I]X = 0$

$$\Rightarrow \begin{bmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

For $\lambda = 0$, we have

$$\begin{bmatrix} 1-0 & 2 \\ 2 & 4-0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \left.\begin{aligned} x_1 + 2x_2 &= 0 \\ 2x_1 + 4x_2 &= 0 \end{aligned}\right\}$$

or $x_1 + 2x_2 = 0$

$$\therefore \frac{x_1}{2} = \frac{-x_2}{1}$$

$$\Rightarrow x_1 = 2, \quad x_2 = -1$$

or $x_1 = 2k, \quad x_2 = -k$

$\therefore X = \begin{bmatrix} 2k \\ -k \end{bmatrix}$ is the eigenvector corresponding to $\lambda = 0$

For $\lambda = 5$, we have

$$\begin{bmatrix} 1-5 & 2 \\ 2 & 4-5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \left.\begin{aligned} -4x_1 + 2x_2 &= 0 \\ 2x_1 - x_2 &= 0 \end{aligned}\right\}$$

or $2x_1 - x_2 = 0$

$$\therefore \frac{x_1}{-1} = \frac{-x_2}{2}$$

$$\Rightarrow \frac{x_1}{1} = \frac{x_2}{2}$$

$$\Rightarrow x_1 = 1, \quad x_2 = 2$$

or $x_1 = k, \quad x_2 = 2k$

$\therefore X = \begin{bmatrix} k \\ 2k \end{bmatrix}$ is the eigenvector corresponding to $\lambda = 5$

3. Find the eigen values and eigen vectors of the matrix $\begin{bmatrix} 1 & 1 & 3 \\ 1 & 5 & 1 \\ 3 & 1 & 1 \end{bmatrix}$

Solution:-

The characteristic equation is $|A-\lambda I|=0$

ie, $\begin{vmatrix} 1-\lambda & 1 & 3 \\ 1 & 5-\lambda & 1 \\ 3 & 1 & 1-\lambda \end{vmatrix} = 0$

$\Rightarrow \quad (1-\lambda)\left[(5-\lambda)(1-\lambda)-1\right] - 1\left[1(1-\lambda)-3\right] + 3\left[1-3(5-\lambda)\right] = 0$

$\Rightarrow \quad (1-\lambda)\left[5-5\lambda-\lambda+\lambda^2-1\right] - \left[1-\lambda-3\right] + 3\left[1-15+3\lambda\right] = 0$

$\Rightarrow \quad (1-\lambda)\left[\lambda^2-6\lambda+4\right] - \left[-\lambda-2\right] + 3\left[3\lambda-14\right] = 0$

$\Rightarrow \quad \lambda^2-6\lambda+4-\lambda^3+6\lambda^2-4\lambda+\lambda+2+9\lambda-42 = 0$

$\Rightarrow \quad -\lambda^3+7\lambda^2-36 = 0$

$\Rightarrow \quad \lambda^3-7\lambda^2+36 = 0$

Solving, $\Rightarrow \quad \lambda = -2, 3, 6$

If $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is an eigen vector corresponding to the eigen value λ, then $[A-\lambda I]X=0$

$\longrightarrow \quad \begin{bmatrix} 1-\lambda & 1 & 3 \\ 1 & 5-\lambda & 1 \\ 3 & 1 & 1-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

For $\lambda = -2$, we have

$$\begin{bmatrix} 3 & 1 & 3 \\ 1 & 7 & 1 \\ 3 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \frac{x_1}{7\times 3 - 1\times 1} = \frac{-x_2}{1\times 3 - 3\times 1} = \frac{x_3}{1\times 1 - 3\times 7}$$

$$\Rightarrow \frac{x_1}{20} = \frac{-x_2}{0} = \frac{x_3}{-20}$$

$$\Rightarrow \frac{x_1}{1} = \frac{x_2}{0} = \frac{x_3}{-1}$$

~~or~~ $\therefore$ $X = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ or $\begin{bmatrix} k \\ 0 \\ -k \end{bmatrix}$ is the eigenvector corresponding to $\lambda = -2$.

For $\lambda = 3$, we have

$$\begin{bmatrix} -2 & 1 & 3 \\ 1 & 2 & 1 \\ 3 & 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \frac{x_1}{-4-1} = \frac{-x_2}{-2-3} = \frac{x_3}{1-6}$$

$$\Rightarrow \frac{x_1}{-5} = \frac{-x_2}{-5} = \frac{x_3}{-5}$$

$$\Rightarrow \frac{x_1}{1} = \frac{x_2}{-1} = \frac{x_3}{1}$$

~~or~~ $\therefore$ $X = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ or $\begin{bmatrix} k \\ -k \\ k \end{bmatrix}$ is the eigenvector corresponding to $\lambda = 3$

For $\lambda = 6$, we have

$$\begin{bmatrix} -5 & 1 & 3 \\ 1 & -1 & 1 \\ 3 & 1 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \frac{x_1}{4} = \frac{-x_2}{-8} = \frac{x_3}{4}.$$

$$\Rightarrow \frac{x_1}{1} = \frac{x_2}{2} = \frac{x_3}{1}$$

~~or~~ $\therefore$ $X = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ or $\begin{bmatrix} k \\ 2k \\ k \end{bmatrix}$ is the eigenvector corresponding to $\lambda = 6$

4. Find the eigen values and eigen vectors of the matrix $\begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 2 \end{bmatrix}$

Solution:

The characteristic equation is $|A-\lambda I| = 0$

i.e. $\begin{vmatrix} 2-\lambda & 1 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 1 & 2-\lambda \end{vmatrix} = 0$

$\Rightarrow (2-\lambda)[(1-\lambda)(2-\lambda)-0] - 1[0-0] + 1[0-(1-\lambda)] = 0$

$\Rightarrow (2-\lambda)(1-\lambda)(2-\lambda) - (1-\lambda) = 0$

$\Rightarrow (1-\lambda)[(2-\lambda)(2-\lambda) - 1] = 0$

$\Rightarrow (1-\lambda)(4-2\lambda-2\lambda+\lambda^2-1) = 0$

$\Rightarrow (1-\lambda)(\lambda^2-4\lambda+3) = 0$

$\Rightarrow (1-\lambda)(\lambda-1)(\lambda-3) = 0$

$\Rightarrow \lambda = 1, 1, 3$ are the eigen values.

If $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is the eigen vector corresponding to the eigen value λ, then $[A-\lambda I]X = 0$

$\Rightarrow \begin{bmatrix} 2-\lambda & 1 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 1 & 2-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

For $\lambda = 1$, we have

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow \quad x_1 + x_2 + x_3 = 0$

$\Rightarrow \quad x_1 = -x_2 - x_3$

let $\quad x_2 = k_1, \; x_3 = k_2$

Then $\quad x_1 = -k_1 - k_2$

$\therefore \quad X = \begin{bmatrix} -k_1 - k_2 \\ k_1 \\ k_2 \end{bmatrix}$ is the eigen vector corresponding to $\lambda = 1$.

For $\lambda = 3$, we have

$$\begin{bmatrix} -1 & 1 & 1 \\ 0 & -2 & 0 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow \quad \frac{x_1}{2} = \frac{-x_2}{0} = \frac{x_3}{2}$

$\Rightarrow \quad \frac{x_1}{1} = \frac{x_2}{0} = \frac{x_3}{1}$

$\therefore \quad X = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ or $\begin{bmatrix} k \\ 0 \\ k \end{bmatrix}$ is the eigen vector corresponding to $\lambda = 3$.

Diagonalization of a matrix

Suppose the n by n matrix A has n linearly independent eigenvectors. If these eigenvectors are the columns of a matrix P, then $P^{-1}AP$ is a diagonal matrix D.

The eigenvalues of A are on the diagonal of D

$$P^{-1}AP = D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

* Any matrix with distinct eigenvalues can be diagonalized.
* The diagonalization matrix P is not unique.
* Not all matrices posses n linearly independent eigenvectors, so not all matrices are diagonalizable
* Diagonalizability of A depends on enough eigenvectors. Invertibility of A depends on nonzero eigenvalues.
* Diagonalization can fail only if there are repeated eigenvalues.
* The eigenvalues of A^k are $\lambda_1^k, \ldots, \lambda_n^k$ and each eigenvector of A is still an eigenvector of A^k.
  $[D^k = (P^{-1}AP)(P^{-1}AP) \cdots (P^{-1}AP) = P^{-1}A^kP]$.
* Diagonalizable matrices share the same eigenvector matrix iff $AB=BA$

example $A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

$|A - \lambda I| = 0 \Rightarrow \lambda^2 - \lambda = 0 \Rightarrow \lambda = 0, 1$.

For $\lambda = 0$ $A - 0I = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \Rightarrow \frac{1}{2}x_1 + \frac{1}{2}x_2 = 0 \Rightarrow x_1 = -x_2$

$\therefore\ X = \begin{bmatrix} k \\ -k \end{bmatrix}$ or $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

For $\lambda = 1$ $A - I = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \Rightarrow -\frac{1}{2}x_1 + \frac{1}{2}x_2 = 0 \Rightarrow x_1 = x_2$

$\therefore\ X = \begin{bmatrix} k \\ k \end{bmatrix}$ or $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\therefore\ P = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ $D = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ $P^{-1} = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix}$

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

$|A - \lambda I| = 0 \Rightarrow \lambda^2 + 1 = 0 \Rightarrow \lambda = -i, +i$

For $\lambda = -i$ $A + iI = \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \Rightarrow ix_1 - x_2 = 0 \Rightarrow x_2 = ix_1$

$\therefore\ X = \begin{bmatrix} k \\ ik \end{bmatrix}$ or $\begin{bmatrix} 1 \\ i \end{bmatrix}$

For $\lambda = i$ $A - iI = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \Rightarrow -ix_1 - x_2 = 0 \Rightarrow x_2 = -ix_1$

$\therefore\ X = \begin{bmatrix} k \\ -ik \end{bmatrix}$ or $\begin{bmatrix} 1 \\ -i \end{bmatrix}$

$\therefore\ P = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$ $D = \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}$ $P^{-1} = \frac{1}{-2i}\begin{bmatrix} -i & -1 \\ -i & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & -i/2 \\ 1/2 & i/2 \end{bmatrix}$

* Diagonalize the matrix if possible.

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

$$|A-\lambda I| = 0 \Rightarrow \lambda^3 - (1-5+1)\lambda^2 + (4-8+4)\lambda - (4-3\times6+3\times6) = 0$$

$$\Rightarrow \lambda^3 + 3\lambda^2 + 0\lambda - 4 = 0 \Rightarrow \lambda = 1, -2, -2$$

For $\lambda = 1$

$$A - I = \begin{bmatrix} 0 & 3 & 3 \\ -3 & -6 & -3 \\ 3 & 3 & 0 \end{bmatrix} \Rightarrow \frac{x_1}{9} = \frac{-x_2}{9} = \frac{x_3}{9} \Rightarrow x = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

For $\lambda = -2$

$$A + 2I = \begin{bmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{bmatrix} \Rightarrow \frac{x_1}{0} = \frac{-x_2}{0} = \frac{x_3}{0}$$

~~eigenvector~~ ~~can't~~ be zero vector.

$3x_1 + 3x_2 + 3x_3 = 0$

$\therefore \Rightarrow x_1 = -x_2 - x_3$

or $x = \begin{bmatrix} -x_2 - x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

$\therefore \quad x = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ and $x = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

$\therefore \quad P = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}$ where $A = PDP^{-1}$.

* Diagonalize the matrix if possible.

$$A = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}.$$

$$|A - \lambda I| = 0 \Rightarrow \lambda^3 + 3\lambda^2 - 0\lambda - 4 = 0 \Rightarrow \lambda = 1, -2, -2$$

For $\lambda = 1$

$$A - I = \begin{bmatrix} 1 & 4 & 3 \\ -4 & -7 & -3 \\ 3 & 3 & 0 \end{bmatrix} \Rightarrow \frac{x_1}{9} = \frac{-x_2}{9} = \frac{x_3}{9} \Rightarrow x = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

For $\lambda = -2$

$$A + 2I = \begin{bmatrix} 4 & 4 & 3 \\ -4 & -4 & -3 \\ 3 & 3 & 3 \end{bmatrix} \Rightarrow \frac{x_1}{-3} = \frac{-x_2}{-3} = \frac{x_3}{0} \Rightarrow x = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

$R_2 + R_1$; $4R_3 - 3R_1$

$$\Rightarrow \begin{bmatrix} 4 & 4 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

since the rank of the echelon form is 2, there exists only 1 free variable, and hence only one eigenvector.

Hence the matrix cannot be diagonalized.

* Diagonalize the matrix if possible.

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

$|A-\lambda I| = 0 \Rightarrow \lambda^3 + 3\lambda^2 + 2\lambda - 0 = 0 \Rightarrow \lambda = -2, -1, 0$

For $\lambda = -2$.

$$A + 2I = \begin{bmatrix} 4 & -2 & -2 \\ 3 & -1 & -2 \\ 2 & -2 & 0 \end{bmatrix} \Rightarrow \frac{x_1}{-4} = \frac{-x_2}{4} = \frac{x_3}{-4} \Rightarrow x = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

For $\lambda = -1$

$$A + I = \begin{bmatrix} 3 & -2 & -2 \\ 3 & -2 & -2 \\ 2 & -2 & -1 \end{bmatrix} \Rightarrow \frac{x_1}{-2} = \frac{-x_2}{1} = \frac{x_3}{-2} \Rightarrow x = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$$

For $\lambda = 0$

$$A + 0I = \begin{bmatrix} 2 & -2 & -2 \\ 3 & -3 & -2 \\ 2 & -2 & -2 \end{bmatrix} \Rightarrow \frac{x_1}{2} = \frac{-x_2}{-2} = \frac{x_3}{0} \Rightarrow x = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\therefore P = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 0 \end{bmatrix} \quad D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad P^{-1} = \begin{bmatrix} -2 & 2 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$$

where $A = PDP^{-1}$

* Diagonalize the matrix if possible.

$$A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

$|A-\lambda I| = 0 \Rightarrow \lambda^3 - 12\lambda^2 + 21\lambda + 98 = 0 \Rightarrow \lambda = -2, 7, 7$

For $\lambda = -2$

$$A + 2I = \begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix} \Rightarrow \frac{x_1}{36} = \frac{-x_2}{-18} = \frac{x_3}{-36} \Rightarrow x = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$$

For $\lambda = 7$

$$A \mp 7I = \begin{bmatrix} -4 & -2 & 4 \\ -2 & -1 & 2 \\ 4 & 2 & -4 \end{bmatrix} \Rightarrow \frac{x_1}{0} = \frac{-x_2}{0} = \frac{x_3}{0}$$

eigen vector can't be zero vector

reducing to echelon form

$2R_2 - R_1$, $R_3 + R_1 \Rightarrow \begin{bmatrix} -4 & -2 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$\therefore -4x_1 - 2x_2 + 4x_3 = 0 \Rightarrow x_1 = -\frac{1}{2}x_2 + x_3 \quad \therefore x = \begin{bmatrix} -\frac{1}{2}x_2 + x_3 \\ x_2 \\ x_3 \end{bmatrix} \therefore x = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, x = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

$$\therefore P = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 2 & 0 \\ -2 & 0 & 1 \end{bmatrix} \quad D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix} \quad P^{-1} = \frac{1}{9}\begin{bmatrix} 2 & 1 & -2 \\ -1 & 4 & 1 \\ 4 & 2 & 5 \end{bmatrix}$$

where $A = PDP^{-1}$.

Singular Value Decomposition

Any m by n matrix A can be factored into

$$A = U \Sigma V^T = (\text{orthogonal})(\text{diagonal})(\text{orthogonal})$$

The columns of U (m by m) are eigenvectors of AA^T, and the columns of V (n by n) are eigenvectors of A^TA. The r singular values on the diagonal of Σ (m by n) are the square roots of the nonzero eigenvalues of both AA^T and A^TA.

Remark

For positive definite matrices, Σ is Λ and $U \Sigma V^T$ is identical to $Q \Lambda Q^T$.

For other symmetric matrices, ~~Σ remains real~~ any negative eigenvalues in Λ become positive in Σ.

Remark.

U and V give orthonormal bases for all four fundamental subspaces:

first r columns of U : column space of A

last $m-r$ columns of U : left nullspace of A

first r columns of V : row space of A

last $n-r$ columns of V : nullspace of A.

The diagonal (but rectangular) matrix Σ has eigenvalues from A^TA. These positive entries (also called sigma) will be $\sigma_1, \ldots, \sigma_r$. They are the singular values of A.

Remark:
When A multiplies a column v_j of V, it produces σ_j times a column of U. ($A = U\Sigma V^T \Rightarrow AV = U\Sigma$).

Remark:
Eigenvectors of AA^T and A^TA must go into the columns of U and V:

$$AA^T = (U\Sigma V^T)(U\Sigma V^T)^T = (U\Sigma V^T)(V\Sigma^T U^T) = U\Sigma\Sigma^T U^T$$

$\therefore$ U must be the eigenvector matrix of AA^T.
The eigenvalue matrix $\Sigma\Sigma^T$ is an $m\times m$ matrix with $\sigma_1^2, \ldots, \sigma_r^2$ on the diagonal.

$$A^TA = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^T U^T U\Sigma V^T = V\Sigma^T\Sigma V^T$$

$\therefore$ V must be the eigenvector matrix of A^TA.
The diagonal matrix $\Sigma^T\Sigma$ has the same $\sigma_1^2, \ldots \sigma_r^2$, but it is $n\times n$.

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \qquad D \rightarrow \underbrace{\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0}_{\text{TUC}}$$

Decompose $A = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$ as $U \Sigma V^T$ where U and V are orthogonal matrices.

Sol^n:

$$A A^T = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}_{3\times1} \begin{bmatrix} -1 & 2 & 2 \end{bmatrix}_{1\times3}$$

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

$$|AA^T - \lambda I| = 0 \Rightarrow \begin{vmatrix} 1-\lambda & -2 & -2 \\ -2 & 4-\lambda & 4 \\ -2 & 4 & 4-\lambda \end{vmatrix} = 0$$

$$(1-\lambda)\left[(4-\lambda)^2 - 16\right] + 2\left[-2(4-\lambda)+8\right] - 2\left[-8 + 2(4-\lambda)\right] = 0$$

$$(1-\lambda)\left[16 - 8\lambda + \lambda^2 - 16\right] + 2\left[\cancel{-8} + 2\lambda \cancel{+8}\right] - 2\left[\cancel{-8} + \cancel{8} - 2\lambda\right] = 0$$

$$(1-\lambda)(\lambda^2 - 8\lambda) + 4\lambda + 4\lambda = 0$$

$$\lambda^2 - \cancel{8\lambda} - \lambda^3 + 8\lambda^2 + \cancel{8\lambda} = 0$$

$$-\lambda^3 + 9\lambda^2 = 0$$

$$-\lambda^2(\lambda - 9) = 0$$

$$\lambda_1 = 0, \ \lambda_2 = 0, \ \lambda_3 = 9$$

$\lambda = 9$

$$(AA^T - \lambda I)X = 0$$

$$\begin{bmatrix} -8 & -2 & -2 \\ -2 & -5 & 4 \\ -2 & 4 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\sim \begin{bmatrix} -8 & -2 & -2 \\ 0 & -18 & 18 \\ 0 & 18 & -18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\Rightarrow \left.\begin{aligned} -8x_1 - 2x_2 - 2x_3 &= 0 \\ -18x_2 + 18x_3 &= 0 \end{aligned}\right\} \quad x_2 = x_3, \quad x_1 = -\tfrac{1}{2}x_3$$

$$X_1 = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$$

$\lambda = 0$

$$(AA^T - \lambda I)X = 0$$

$$\begin{bmatrix} 1 & -2 & -2 \\ -2 & 4 & 4 \\ -2 & 4 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\sim \begin{bmatrix} 1 & -2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\Rightarrow x_1 - 2x_2 - 2x_3 = 0$$

$$\Rightarrow x_1 = 2x_2 + 2x_3$$

let $\cancel{x_2 = -\tfrac{1}{3}}$ $x_2 = -1, \ x_3 = 2$ $\qquad X_2 = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$

let $x_2 = 2, \ x_3 = -1$ $\qquad X_3 = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$

$$\therefore U = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{bmatrix}$$

$$A^T A = \begin{bmatrix} -1 & 2 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 9 \end{bmatrix}$$

$$|A^T A - \lambda I| = 0 \Rightarrow |9 - \lambda| = 0 \Rightarrow \underline{\lambda = 9}$$

Then $(A^T A - \lambda I) X = 0$

$\Rightarrow \quad 0 . x_1 = 0$

let $x_1 = 1$ $\quad \therefore X = [1]$

$\therefore V = [1]$ or $V^T = [1]$.

9 is an eigenvalue of both AA^T and $A^T A$.

And rank of $A = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$ is $r = 1$.

$\therefore \Sigma$ has only $\sigma_1 = \sqrt{9} = 3$. $\quad \therefore \Sigma = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$

$\therefore$ The SVD of $A = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix}$

Obtain the SVD of $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ $\begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$

Sol^n

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

$|AA^T - \lambda I| = 0 \Rightarrow \begin{vmatrix} 2-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix} = 0 \Rightarrow (2-\lambda)(1-\lambda) - 1 = 0 \Rightarrow \lambda^2 - 3\lambda + 1 = 0$

$\Rightarrow \lambda = \dfrac{3 \pm \sqrt{5}}{2}$

$\lambda_1 = \dfrac{3 - \sqrt{5}}{2}$

$(AA^T - \lambda I)X = 0$

$\Rightarrow \begin{bmatrix} 2 - \left(\frac{3-\sqrt{5}}{2}\right) & 1 \\ 1 & 1 - \left(\frac{3-\sqrt{5}}{2}\right) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$

$\begin{bmatrix} \frac{1+\sqrt{5}}{2} & 1 \\ 1 & \frac{-1+\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$

$\Rightarrow \dfrac{1+\sqrt{5}}{2} x_1 + x_2 = 0$

$\Rightarrow x_1 = -\dfrac{2}{1+\sqrt{5}} x_2$

let $x_2 = \dfrac{1+\sqrt{5}}{2} \Rightarrow x_1 = -1$

$X = \begin{bmatrix} -1 \\ \frac{1+\sqrt{5}}{2} \end{bmatrix} = \begin{bmatrix} -1 \\ \alpha \end{bmatrix}$

$\alpha = \dfrac{1+\sqrt{5}}{2}$

$\lambda_2 = \dfrac{3+\sqrt{5}}{2}$

$(AA^T - \lambda I)X = 0$

$\Rightarrow \begin{bmatrix} 2 - \left(\frac{3+\sqrt{5}}{2}\right) & 1 \\ 1 & 1 - \left(\frac{3+\sqrt{5}}{2}\right) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$

$\begin{bmatrix} \frac{1-\sqrt{5}}{2} & 1 \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$

$\Rightarrow \dfrac{1-\sqrt{5}}{2} x_1 + x_2 = 0$

$\Rightarrow x_1 = -\dfrac{2}{1-\sqrt{5}} x_2$

let $x_2 = \dfrac{1-\sqrt{5}}{2} \Rightarrow x_1 = -1$

$X = \begin{bmatrix} -1 \\ \frac{1-\sqrt{5}}{2} \end{bmatrix} = \begin{bmatrix} -1 \\ \beta \end{bmatrix}$

$\beta = \dfrac{1-\sqrt{5}}{2}$

$$U = \begin{bmatrix} \frac{-1}{\sqrt{1+\alpha^2}} & \frac{-1}{\sqrt{1+\beta^2}} \\ \frac{\alpha}{\sqrt{1+\alpha^2}} & \frac{\beta}{\sqrt{1+\beta^2}} \end{bmatrix}$$

As $A^TA = AA^T$ $\quad V^T = \begin{bmatrix} \frac{-1}{\sqrt{1+\alpha^2}} & \frac{\alpha}{\sqrt{1+\alpha^2}} \\ \frac{-1}{\sqrt{1+\beta^2}} & \frac{\beta}{\sqrt{1+\beta^2}} \end{bmatrix}$ and $\Sigma = \begin{bmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{bmatrix}$

Obtain the SVD of $A = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}_{2\times 3}$

$$AA^T = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}_{2\times 3} \begin{bmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}_{3\times 2} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$A^TA = \begin{bmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}_{3\times 2} \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}_{2\times 3} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

$|AA^T - \lambda I| = 0$

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

$\Rightarrow (2-\lambda)^2 - 1 = 0$

$\lambda^2 - 4\lambda + 3 = 0$

$(\lambda-1)(\lambda-3) = 0$

$\lambda = 1, \quad \lambda = 3$

$\underline{\lambda_1 = 3}$ $\quad (AA^T - \lambda I)X = 0$

$\begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$

$x_1 = -x_2$

Let $x_2 = 1, x_1 = -1$

$x = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

$\underline{\lambda_2 = 1}$

$(AA^T - \lambda I)X = 0$

$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$

$\Rightarrow x_1 = x_2$

Let $x_2 = 1, x_1 = 1$

$x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$$U = \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}_{2\times 2}$$

$$\Sigma = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$|A^TA - \lambda I| = 0$

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

$\Rightarrow (1-\lambda)\left[(2-\lambda)(1-\lambda) - 1\right] + 1\left[-(1-\lambda)\right] = 0$

$(1-\lambda)(2 - 2\lambda - \lambda + \lambda^2 - 1) - 1 + \lambda = 0$

$(1-\lambda)(2 - 2\lambda - \lambda + \lambda^2 - 1) - 1 + \lambda = 0$

$(1-\lambda)(\lambda^2 - 3\lambda + 1) - 1 + \lambda = 0$

$(1-\lambda)(\lambda^2 - 3\lambda + 1) \quad \lambda^2 - 3\lambda + \cancel{1} - \lambda^3 + 3\lambda^2 - \cancel{\lambda} - \cancel{1} + \cancel{\lambda} = 0$

$-\lambda^3 + 4\lambda^2 - 3\lambda = 0$

$-\lambda(\lambda^2 - 4\lambda + 3) = 0$

$\lambda_1 = 0, \ \lambda_2 = 1, \ \lambda_3 = 3$

$\underline{\lambda_1 = 0}$

$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$

$\sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$

$x_1 - x_2 = 0 \qquad x_1 = x_2$

$x_2 - x_3 = 0 \qquad x_2 = x_3$

$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

$\underline{\lambda = 1}$

$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$

$\sim \begin{bmatrix} 0 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$

$-x_2 = 0 \qquad x_2 = 0$

$-x_1 + x_2 - x_3 = 0 \qquad x_1 = -x_3$

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

$\underline{\lambda = 3}$

$\begin{bmatrix} -2 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & -1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$

$\sim \begin{bmatrix} -2 & -1 & 0 \\ 0 & 1 & 2 \\ 0 & -1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$

$-2x_1 - x_2 = 0$

$x_2 + 2x_3 = 0$

$x_2 = -2x_1$

$\begin{bmatrix} 1 \\ -2 \\ +1 \end{bmatrix}$

$$V = \begin{bmatrix} 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ -2/\sqrt{6} & 0 & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix} \qquad V^T = \begin{bmatrix} 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}_{3\times 3}$$