Course Title: Linear Algebra, Laplace Transforms and Combinatorics
Course Code: 18MA31A
Unit 2: Linear Algebra II

Department of Mathematics
RV College of Engineering
Bangalore 560059

Topic Learning Objectives:

- Study the orthogonal and orthonormal properties of vectors.
- Study orthogonal projections.
- Use Gram-Schmidt process to factorize a given matrix as a product of an orthogonal matrix(Q) and an upper triangular invertible matrix(R).
- Diagonalize symmetric matrices using eigenvalues and eigenvectors. $D = P^{-1}AP$
- Decompose a given matrix into product of an orthogonal matrix(U), a diagonal matrix (Σ) and an orthogonal matrix(V^T).

Orthogonality

- u and $v \in \mathbb{R}^n$ - orthogonal if $u \cdot v = 0$
- $u = (1, 2)$, $v = (6, -3) \in \mathbb{R}^2$ - orthogonal $\because$ $u \cdot v = 0$
- $\{u_1, u_2, ..., u_p\}$ in $\mathbb{R}^n$ - orthogonal set if $u_i.u_j = 0$ whenever $i \neq j$
  ex.$\{u_1, u_2, u_3\}$ such that $u_1 = (3, 1, 1)$, $u_2 = (-1, 2, 1)$, $u_3 = (-\frac{1}{2}, -2, \frac{7}{2})$ is an orthogonal set, since $u_1.u_2 = 0$, $u_1.u_3 = 0$, $u_2.u_3 = 0$.
- A set $\{u_1, u_2, ..., u_p\}$ - orthonormal set if $u_i.u_j = 0$ whenever $i \neq j$, $u_i.u_i = 1$ whenever $i = j$
  ex., $\{e_1, e_2, ..., e_n\}$ - the standard basis for $\mathbb{R}^n$, is an orthonormal set

- 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 $\mathbb{R}^3$
- An orthonormal basis for a subspace W of $\mathbb{R}^n$ is a basis for W that is also an orthonormal set.
  ex. $\{v_1, v_2, v_3\}$, where
  $v_1 = (\frac{3}{\sqrt{11}}, \frac{1}{\sqrt{11}}, \frac{1}{\sqrt{11}}), v_2 = (-\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}), v_3 = (-\frac{1}{\sqrt{66}}, -\frac{4}{\sqrt{66}}, \frac{7}{\sqrt{66}})$ is an orthonormal basis of $\mathbb{R}^3$

- A square matrix A satisfying the condition $A^{-1} = A^T$ - orthogonal matrix.

  ex. $P = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

  then $P^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$ and $P^T = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$

  $\therefore$ P is an orthogonal 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

- The row vectors of A are orthonormal

  similarly the column vectors

Orthogonal Projections:

Given a non-zero vector $\overrightarrow{u}$ in $\mathbb{R}^n$, consider the problem of decomposing a vector $\overrightarrow{y}$ in $\mathbb{R}^n$ into the sum of two vectors, one a multiple of $\overrightarrow{u}$ and the other orthogonal to $\overrightarrow{u}$.
We wish to write $\overrightarrow{y} = \hat{y} + \overrightarrow{z} - (1)$, where $\hat{y} = \alpha \overrightarrow{u}$, for some scalar α and $\overrightarrow{z}$ is some vector orthogonal to $\overrightarrow{u}$.

Given any scalar α, let $\overrightarrow{z} = \overrightarrow{y} - \alpha \overrightarrow{u}$, so that (1) is satisfied.
Then $\overrightarrow{y} - \hat{y}$ is orthogonal to $\overrightarrow{u}$ iff

$$0 = (\overrightarrow{y} - \alpha \overrightarrow{u}).\overrightarrow{u} = \overrightarrow{y}.\overrightarrow{u} - (\alpha \overrightarrow{u}).\overrightarrow{u} = \overrightarrow{y}.\overrightarrow{u} - \alpha(\overrightarrow{u}.\overrightarrow{u})$$

That is, (1) is satisfied with $\overrightarrow{z}$ orthogonal to $\overrightarrow{u}$ iff

$$\alpha = \frac{\overrightarrow{y}.\overrightarrow{u}}{\overrightarrow{u}.\overrightarrow{u}} \quad \text{and} \quad \hat{y} = \frac{\overrightarrow{y}.\overrightarrow{u}}{\overrightarrow{u}.\overrightarrow{u}}\overrightarrow{u}$$

The vector $\hat{y}$ is called the orthogonal projection of $\overrightarrow{y}$ onto $\overrightarrow{u}$, and the vector $\overrightarrow{z}$ is called the component of $\overrightarrow{y}$ orthogonal to $\overrightarrow{u}$.

Orthogonal Projections:

Given a non-zero vector $\overrightarrow{u}$ in $\mathbb{R}^n$, consider the problem of decomposing a vector $\overrightarrow{y}$ in $\mathbb{R}^n$ into the sum of two vectors, one a multiple of $\overrightarrow{u}$ and the other orthogonal to $\overrightarrow{u}$.

We wish to write $\overrightarrow{y} = \hat{y} + \overrightarrow{z} - (1)$, where $\hat{y} = \alpha \overrightarrow{u}$, for some scalar α and $\overrightarrow{z}$ is some vector orthogonal to $\overrightarrow{u}$.

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

Then $\overrightarrow{y} - \hat{y}$ is orthogonal to $\overrightarrow{u}$ iff

$$0 = (\overrightarrow{y} - \alpha \overrightarrow{u}).\overrightarrow{u} = \overrightarrow{y}.\overrightarrow{u} - (\alpha \overrightarrow{u}).\overrightarrow{u} = \overrightarrow{y}.\overrightarrow{u} - \alpha(\overrightarrow{u}.\overrightarrow{u})$$

That is, (1) is satisfied with $\overrightarrow{z}$ orthogonal to $\overrightarrow{u}$ iff

$$\alpha = \frac{\overrightarrow{y}.\overrightarrow{u}}{\overrightarrow{u}.\overrightarrow{u}} \quad \text{and} \quad \hat{y} = \frac{\overrightarrow{y}.\overrightarrow{u}}{\overrightarrow{u}.\overrightarrow{u}}\overrightarrow{u}$$

The vector $\hat{y}$ is called the orthogonal projection of $\overrightarrow{y}$ onto $\overrightarrow{u}$, and the vector $\overrightarrow{z}$ is called the component of $\overrightarrow{y}$ orthogonal to $\overrightarrow{u}$.

Example

Let $\overrightarrow{y} = (7,6)$ and $\overrightarrow{u} = (4,2)$.
The orthogonal projection of $\overrightarrow{y}$ onto $\overrightarrow{u}$ is given by,

$$\hat{y} = \frac{\overrightarrow{y}\,.\,\overrightarrow{u}}{\overrightarrow{u}\,.\,\overrightarrow{u}}\overrightarrow{u} = \frac{40}{20}\overrightarrow{u} = 2\overrightarrow{u} = 2(4,2) = (8,4)$$

$z = y - \hat{y} = (-1,2)$

Orthogonal projection

The orthogonal projection of $\overrightarrow{y}$ onto a space W spanned by orthogonal vectors $\{u_1, u_2\}$ is given by $\hat{y} = \dfrac{\overrightarrow{y}.\overrightarrow{u_1}}{\overrightarrow{u_1}.\overrightarrow{u_1}}\overrightarrow{u_1} + \dfrac{\overrightarrow{y}.\overrightarrow{u_2}}{\overrightarrow{u_2}.\overrightarrow{u_2}}\overrightarrow{u_2}$

The distance from a point $\overrightarrow{y}$ in $\mathbb{R}^n$ to a subspace W is defined as the distance from $\overrightarrow{y}$ to the nearest point in W.

Example

The distance from $\overrightarrow{y}$ to $W = \text{Span}\{u_1, u_2\}$, where $\overrightarrow{y} = (-1, -5, 10), u_1 = (5, -2, 1), u_2 = (1, 2, -1)$. is given by

$$\hat{y} = \frac{(-1, -5, 10).(5, -2, 1)}{(5, -2, 1).(5, -2, 1)}(5, -2, 1) + \frac{(-1, -5, 10).(1, 2, -1)}{(1, 2, -1).(1, 2, -1)}(1, 2, -1)$$

$$= (-1, -8, 4)$$

$$\overrightarrow{y} - \hat{y} = (-1, -5, 10) - (-1, -8, 4) = (0, 3, 6)$$

The distance from $\overrightarrow{y}$ to W is $\sqrt{0 + 3^2 + 6^2} = \sqrt{45} = 3\sqrt{5}$.

Gram-Schmidt Orthogonalization

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

The construction converts a skewed set of axes into a perpendicular set.

Gram-Schmidt process

Given a basis $\{x_1, x_2, ..., x_p\}$ for a subspace W of $\mathbb{R}^n$

define, $v_1 = x_1$

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

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

.

.

.

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

Then $\{v_1, v_2, ..., v_p\}$ is an orthogonal basis for W.

In addition $\mathrm{Span}\{v_1, v_2, ..., v_p\} = \mathrm{Span}\{x_1, x_2, ..., x_k\}$ for $1 \leq k \leq p$.

Example

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.

Solution:

Let $v_1 = x_1 = (3, 6, 0)$

and $v_2 = x_2 - \dfrac{x_2.x_1}{x_1.x_1}x_1 = (1, 2, 2) - \dfrac{(1,2,2).(3,6,0)}{(3,6,0).(3,6,0)}(3, 6, 0) = (0, 0, 2)$.

Then $\{v_1, v_2\}$ is an orthogonal set of non-zero vectors in W.

Since $\dim W = 2$, the set $\{v_1, v_2\}$ is a basis in W.

Example

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.

Solution:

Set $u_1 = v_1 = (0,1,2)$

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

and $u_3 = v_3 - \dfrac{v_3.u_1}{u_1.u_1}u_1 - \dfrac{v_3.u_2}{u_2, u_2}u_2$

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

$= (1,0,1) - \dfrac{2}{5}(0,1,2) - (1,0,0) = (0, -\dfrac{2}{5}, \dfrac{1}{5})$.

QR Factorization:

If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as $A = QR$, where Q is an $m \times n$ matrix whose columns form an orthonormal basis for col A and R is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.

Example

Find a QR factorization of $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

Solution: Construction an orthonormal basis for Col A

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

Let $v_1 = x_1 = (1, 1, 1, 1)$ $\quad v_2 = x_2 - \frac{x_2 . v_1}{v_1 . v_1} v_1 = (-3/4, 1/4, 1/4, 1/4)$

$$v_3 = x_3 - \frac{x_3 . v_1}{v_1 . v_1} v_1 - \frac{x_3 . v_2}{v_2 . v_2} v_2 = (0, -2/3, 1/3, 2/3)$$

$\therefore \{v_1, v_2, v_3\}$ forms an orthogonal basis of Col A .

$$\{(1/2, 1/2, 1/2, 1/2), (-3/\sqrt{12}, 1/\sqrt{12}, 1/\sqrt{12}, 1/\sqrt{12}), (0, -2/\sqrt{6}, 1/\sqrt{6}, 1/\sqrt{6})\}$$

$$\therefore Q = \begin{bmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{bmatrix}$$

<u>To construct an upper triangular invertible matrix</u>

We have $A = QR \implies Q^T A = Q^T QR \implies Q^T A = IR \implies Q^T A = R$ i.e., $R = Q^T A$.

$$\therefore R = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ -3/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} \\ 0 & -2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 2/\sqrt{6} \end{bmatrix}.$$

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,

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

is called the characteristic equation of A.

On expanding the determinant, the characteristic equation takes the form

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

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.

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

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,

$(a_{11} - \lambda)x_1 + a_{12}x_2 + ... + a_{1n}x_n = 0$

$a_{21}x_1 + (a_{22} - \lambda)x_2 + ... + a_{2n}x_n = 0$

......

......

$a_{n1}x_1 + a_{n2}x_2 + ... + (a_{nn} - \lambda)x_n = 0$ ———(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, $x = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{bmatrix}$, 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 ia 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.

Example

Find the Eigen Values and Eigen vectors of the matrix $A = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$

Solution:

$$|A - \lambda I| = 0 \implies \begin{vmatrix} \frac{1}{2} - \lambda & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} - \lambda \end{vmatrix} = 0 \implies \lambda^2 - \lambda = 0 \implies \lambda = 0, \lambda = 1.$$

with $\lambda = 1$, $(A - \lambda I)x = 0 \implies \begin{bmatrix} -1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\implies -x_1 + x_2 = 0 \implies x_2 = x_1$

Letting $x_1 = 1 \implies x_2 = 1 \therefore x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

with $\lambda = 0$, $(A - \lambda I)x = 0 \implies \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\implies x_1 + x_2 = 0 \implies x_2 = -x_1$

Letting $x_1 = 1 \implies x_2 = -1 \therefore x = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Example

Find the Eigen Values and Eigen vectors of the matrix $A = \begin{bmatrix} 3 & 4 & 2 \\ 3 & 5 & 4 \\ 0 & 1 & 2 \end{bmatrix}$.

Solution: $|A - \lambda I| = 0 \implies \begin{vmatrix} 3-\lambda & 4 & 2 \\ 3 & 5-\lambda & 4 \\ 0 & 1 & 2-\lambda \end{vmatrix} = 0 \implies \lambda^3 - 10\lambda^2 + 15\lambda = 0$

$\implies \lambda = 5 + \sqrt{10},\ 5 - \sqrt{10},\ 0.$

Finding Eigen vectors corresponding to each Eigen value

with $\lambda_1 = 5 + \sqrt{10}$, $|A - \lambda I| = 0$

$$\implies \begin{bmatrix} -2-\sqrt{10} & 4 & 2 \\ 3 & -\sqrt{10} & 4 \\ 0 & 1 & -3-\sqrt{10} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\implies \frac{x_1}{3\sqrt{10}+6} = \frac{-x_2}{-9-3\sqrt{10}} = \frac{x_3}{3} \implies \frac{x_1}{2+\sqrt{10}} = \frac{x_2}{3+\sqrt{10}} = \frac{x_3}{1}.$$

$$\therefore x = \begin{bmatrix} 2+\sqrt{10} \\ 3+\sqrt{10} \\ 1 \end{bmatrix}$$

with $\lambda_2 = 5 - \sqrt{10}$, $|A - \lambda I| = 0$

$$\implies \begin{bmatrix} -2+\sqrt{10} & 4 & 2 \\ 3 & \sqrt{10} & 4 \\ 0 & 1 & -3+\sqrt{10} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\implies \frac{x_1}{-3\sqrt{10}+6} = \frac{-x_2}{-9+3\sqrt{10}} = \frac{x_3}{3} \implies \frac{x_1}{2-\sqrt{10}} = \frac{x_2}{3-\sqrt{10}} = \frac{x_3}{1}$$

$$\therefore x = \begin{bmatrix} 2-\sqrt{10} \\ 3-\sqrt{10} \\ 1 \end{bmatrix}$$

with $\lambda_3 = 0$, $|A - \lambda I| = 0$.

$$\implies \begin{bmatrix} 3 & 4 & 2 \\ 3 & 5 & 4 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

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

$$\therefore x = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}.$$

Diagonalization of a Matrix:

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

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

NOTE:

1. Any matrix with distinct eigen values can be diagonalized.
2. The diagonalization matrix P is not unique.
3. Not all matrices posses n linearly independent eigen vectors, so not all matrices are diagonalizable.
4. Diagonalizability of A depends on enough eigen vectors.
5. Diagonalizability can fail only if there are repeated eigen values.
6. The eigen values of A^k are $\lambda_1^k, \lambda_2^k, ..., \lambda_n^k$ and each eigen vector of A is still an eigen vector of A^k.
   $[D^k = D.D....D(k\,\text{times}) = (P^{-1}AP)(P^{-1}AP)...(P^{-1}AP) = P^{-1}A^kP]$.

Example

Diagonalize the matrix $A = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix}$.

Solution: $|A - \lambda I| = 0 \implies \begin{vmatrix} 7-\lambda & 2 \\ 2 & 4-\lambda \end{vmatrix} = 0 \implies \lambda^2 - 11\lambda + 24 = 0$

$\implies (\lambda - 3)(\lambda - 8) = 0 \implies \lambda = 3, \lambda = 8.$

With $\lambda = 3, (A - 3I)x = 0 \implies \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies 2x_1 + x_2 = 0.$

Letting $x_1 = 1 \implies x_2 = -2.$

Hence $\mathrm{x}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$

With $\lambda = 8, (A - 8I)x = 0 \implies \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies -x_1 + 2x_2 = 0.$

Letting $x_2 = 1 \implies x_1 = 2.$

Hence $\mathrm{x}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

$\implies P = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}, D = \begin{bmatrix} 3 & 0 \\ 0 & 8 \end{bmatrix}, P^{-1} = \begin{bmatrix} 1/5 & -2/5 \\ 2/5 & 1/5 \end{bmatrix}$

Example

Diagonalize the matrix $A = \begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$

Solution: $|A - \lambda I| = 0$

$$\implies \begin{vmatrix} 6-\lambda & -2 & -1 \\ -2 & 6-\lambda & -1 \\ -1 & -1 & 5-\lambda \end{vmatrix} = 0$$

$$\implies \lambda^3 - 17\lambda^2 + 90\lambda - 144 = 0$$

$$\implies \lambda = 3,\ 6,\ 8$$

with $\lambda_1 = 3$, $|A - \lambda I| = 0 \implies \begin{bmatrix} 3 & -2 & -1 \\ -2 & 3 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

$\implies \frac{x_1}{5} = \frac{-x_2}{-5} = \frac{x_3}{5} \implies \frac{x_1}{1} = \frac{x_2}{1} = \frac{x_3}{1}$

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

with $\lambda_2 = 6$ $|A - \lambda I| = 0 \implies \begin{bmatrix} 0 & -2 & -1 \\ -2 & 0 & -1 \\ -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

$\implies \dfrac{x_1}{-1} = \dfrac{-x_2}{1} = \dfrac{x_3}{2} \implies \dfrac{x_1}{-1} = \dfrac{x_2}{-1} = \dfrac{x_3}{2}$

$\therefore x = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$

$$\text{with } \lambda_3 = 8 \; |A - \lambda I| = 0 \implies \begin{bmatrix} -2 & -2 & -1 \\ -2 & -2 & -1 \\ -1 & -1 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\implies \frac{x_1}{5} = \frac{-x_2}{5} = \frac{x_3}{0} \implies \frac{x_1}{1} = \frac{x_2}{-1} = \frac{x_3}{2}$$

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

$$\text{Hence } P = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & -1 \\ 1 & 2 & 0 \end{bmatrix}, \; D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 8 \end{bmatrix}, \; P^{-1} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ -1/6 & -1/6 & 1/3 \\ 1/2 & -1/2 & 0 \end{bmatrix}.$$

Singular Value Decomposition:

Any $m \times n$ matrix A can be factored into
$A = U\ \Sigma\ V^T$ = (orthogonal) (diagonal) (orthogonal).
The columns of $U(m \text{ by } m)$ are eigen vectors of AA^T,
and the columns of $V(n \text{ by } n)$ are eigen vectors of $A^T A$.
The r singular values on the diagonal of $\Sigma(m \text{ by } n)$ are the square roots of the non-zero eigen values of both AA^T and $A^T A$.
Note: The singular values are always positive. These positive entries(also called sigma) will be $\sigma_1, \sigma_2, ..., \sigma_r$, such that $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r > 0$.
When A multiplies a column v_j of V, it produces σ_j times a column of U.
$(A = U\Sigma V^T \implies AV = U\Sigma)$.

Example

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

Solution: Consider $AA^T = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} \begin{bmatrix} -1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -2 & -2 \\ -2 & 4 & 4 \\ -2 & 4 & 4 \end{bmatrix}$

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

$\implies \lambda^3 - 9\lambda^2 = 0 \implies \lambda_1 = 0, \lambda_2 = 0, \lambda_3 = 9$

with $\lambda_1 = 9,\ [AA^T - \lambda_1 I]x = 0 \implies \begin{bmatrix} -8 & -2 & -2 \\ -2 & -5 & 4 \\ -2 & 4 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

$\implies -8x_1 - 2x_2 - 2x_3 = 0, -18x_2 + 18x_3 = 0$

$\implies x_1 = -(1/2)x_3,\ x_2 = x_3 \implies x = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$

with $\lambda = \lambda_2 = \lambda_3 = 0, [AA^T - \lambda I]x = 0$

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

Hence $U = \begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix}$

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

$|A^T A - \lambda I| = 0 \implies |9 - \lambda| = 0 \implies \lambda = 9$

Then $(A^T A - \lambda I)x = 0 \implies \begin{bmatrix} 0 \end{bmatrix} \begin{bmatrix} x_1 \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix}$

Let $x_1 = 1 \therefore x = \begin{bmatrix} 1 \end{bmatrix}$

Hence $V = \begin{bmatrix} 1 \end{bmatrix}$ or $V^T = \begin{bmatrix} 1 \end{bmatrix}$.

9 is an eigen value of both AA^T and A^TA.

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

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

$\therefore$ the SVD of $A = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix}$.

Example

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

Solution: Consider, $AA^T = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$

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

$$\implies \lambda^2 - 3\lambda + 1 = 0 \implies \lambda_1 = \frac{3+\sqrt{5}}{2}, \lambda_2 = \frac{3-\sqrt{5}}{2}.$$

with $\lambda_1 = \dfrac{3+\sqrt{5}}{2}$, $(AA^T - \lambda_2 I)x = 0 \implies \begin{bmatrix} \dfrac{1-\sqrt{5}}{2} & 1 \\ 1 & \dfrac{-1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

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

Letting $x_1 = -1$, then $x_2 = \dfrac{1-\sqrt{5}}{2}$ $\therefore x = \begin{bmatrix} -1 \\ \dfrac{1-\sqrt{5}}{2} \end{bmatrix} = \begin{bmatrix} -1 \\ \alpha \end{bmatrix}$, where $\alpha = \dfrac{1-\sqrt{5}}{2}$.

with $\lambda_2 = \dfrac{3-\sqrt{5}}{2}$, $(AA^T - \lambda_1 I)x = 0 \implies \begin{bmatrix} \dfrac{1+\sqrt{5}}{2} & 1 \\ 1 & \dfrac{-1+\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

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

Letting $x_1 = -1$, then $x_2 = \dfrac{1+\sqrt{5}}{2}$ $\therefore x = \begin{bmatrix} -1 \\ \dfrac{1+\sqrt{5}}{2} \end{bmatrix} = \begin{bmatrix} -1 \\ \beta \end{bmatrix}$, where $\beta = \dfrac{1+\sqrt{5}}{2}$.

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

Now $A^T A = \begin{bmatrix} 2 & 1 \\ \\ 1 & 1 \end{bmatrix}$.

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

Hence we have obtained SVD of $A = U\Sigma V^T$.