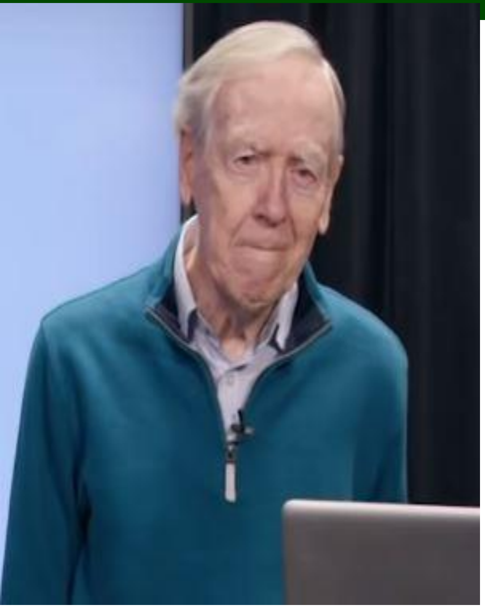
23MAT204

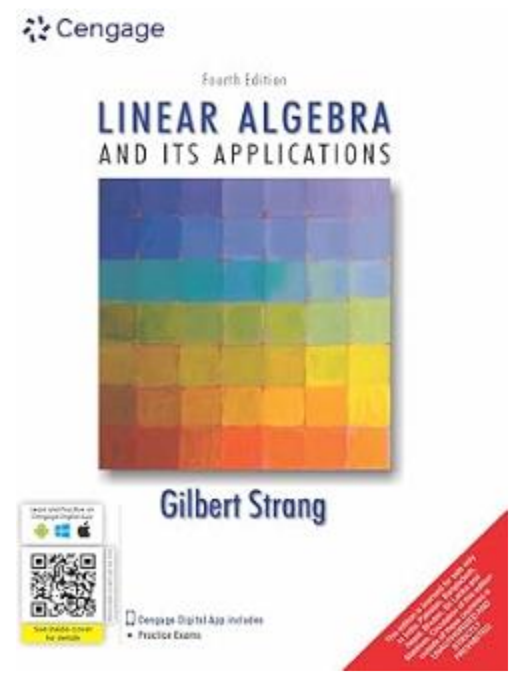
MATHEMATICS FOR INTELLIGENT SYSTEMS-3

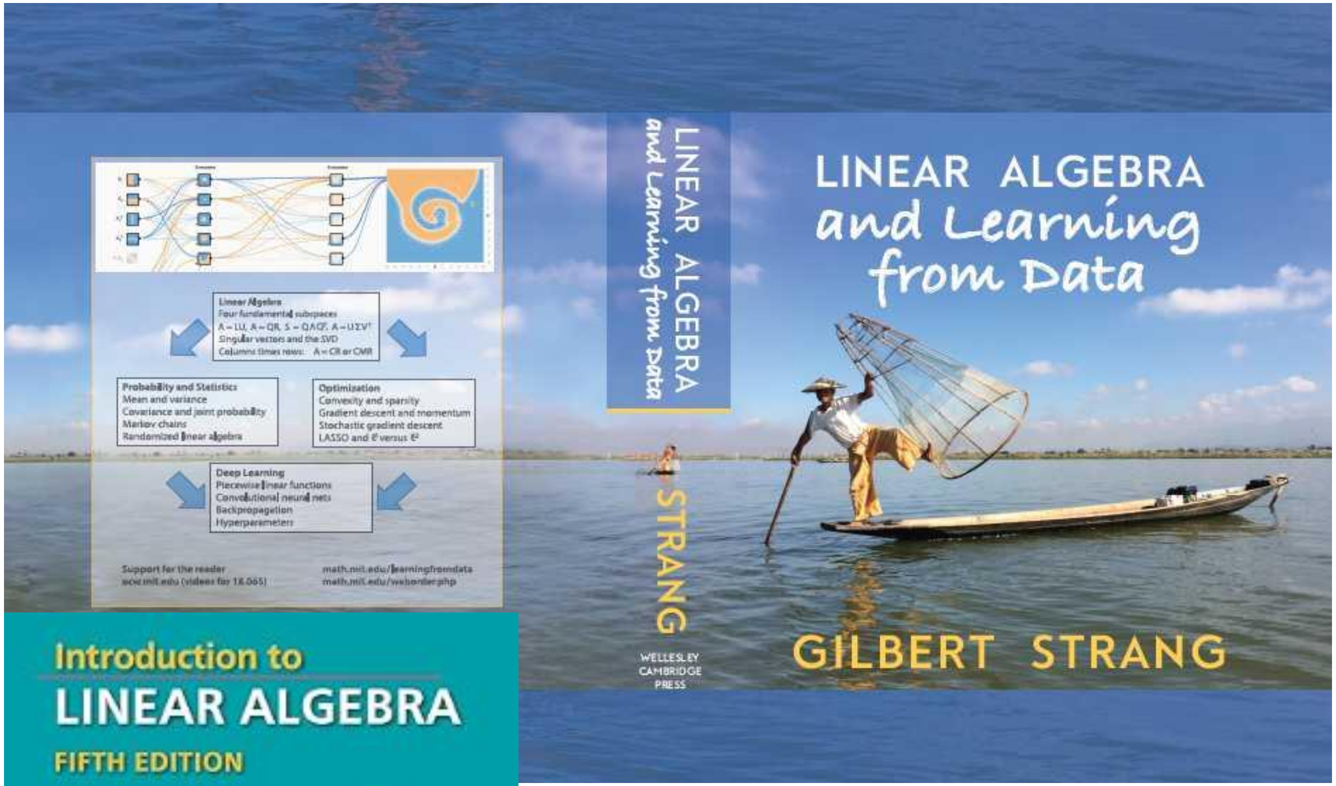
Dr. Sarada Jayan
9448455760 (WhatsApp)
6364333760

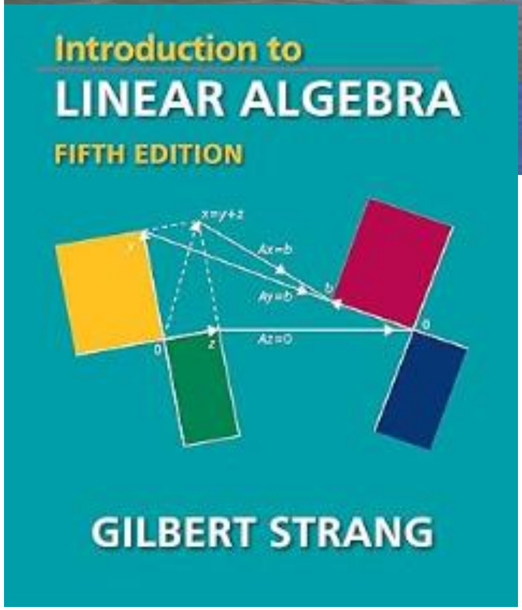
What you learned in MIS-1 and MIS2 in Linear Algebra

- Vectors
- Concept of Linear Independence
- Gaussian Elimination to solve a linear system, AX=B and AX=0
- Parametric Representation of lines, curves
- Vector Spaces
- Subspaces
- Span, Basis and Dimension
- Four fundamental subspaces generated from a matrix
- Orthogonality and orthogonal subspaces
- Eigenvalues and Eigenvectors, their properties
- Matrix Decompositions (CR, LU, QR, SVD, QS)
- Singular Value Decomposition and its uses
- Principal Component Analysis and its uses




Cengage
Fourth Edition
LINEAR ALGEBRA
AND ITS APPLICATIONS
Gilbert Strang
Cengage Digital App includes
Practice Exams


LINEAR ALGEBRA
and Learning from Data
GILBERT STRANG
LINEAR ALGEBRA and Learning from Data
STRANG
WELLESLEY CAMBRIDGE PRESS
Linear Algebra
Four fundamental subspaces
Singular vectors and the SVD
Probability and Statistics
Mean and variance
Covariance and joint probability
Markov chains
Randomized linear algebra
Optimization
Convexity and sparsity
Gradient descent and momentum
Stochastic gradient descent
Deep Learning
Piecewise linear functions
Convolutional neural nets
Backpropagation
Hyperparameters
Support for the reader


Introduction to
LINEAR ALGEBRA
FIFTH EDITION
Ax=b
Ay=b
Az=0
GILBERT STRANG

Vector space definition

First level Layman definition:

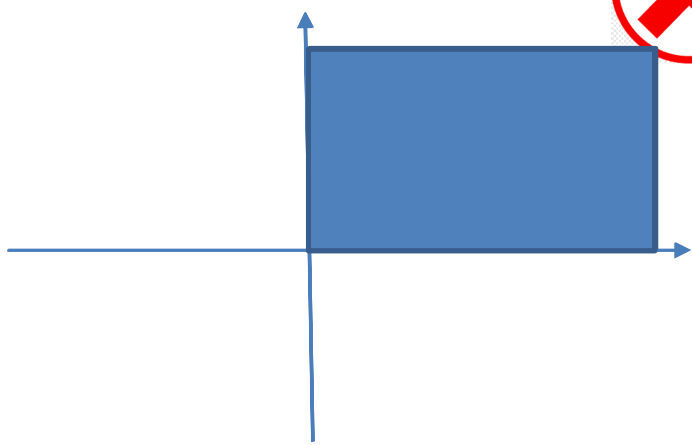
A set of infinite vectors of the same kind (same tuple size) with closure property w.r.t linear combination

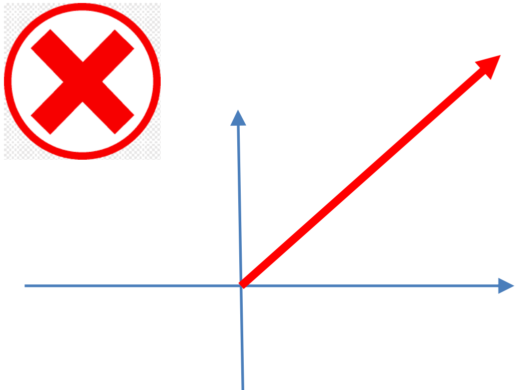
Let V be a vector space. Then v_1, v_2 element of V implies $c_1v_1 + c_2v_2$ is also element of V, where c_1 and c_2 are scalars

Implication:

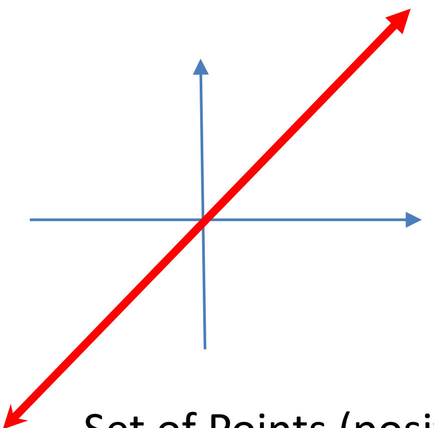
- Zero vector is a necessary element of V
- If v is element of V then –v is also element of V

Set of Points (position vectors) in first quadrant in R^2



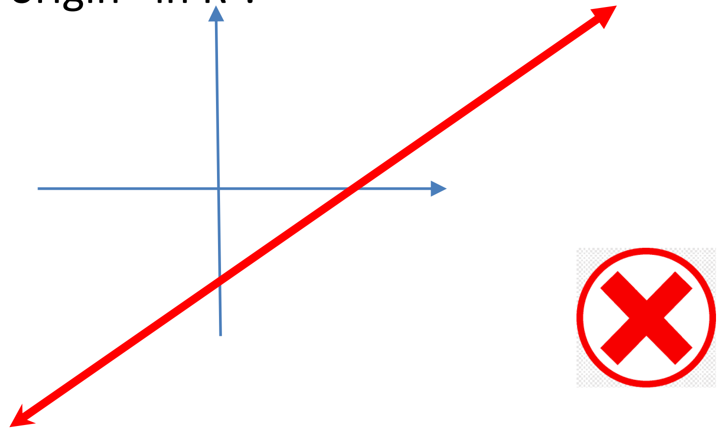


Set of Points(position vectors) on a ray through origin in R^2.



Set of Points (position vectors) on a line through origin in R^2.

Set of Points(position vectors) on a line not through origin in R^2.



Vector spaces

Classical example which we always encounter

$R, R^2, R^3, \ldots. R^{1000}, \ldots\ldots, R^{\text{infinity}}$

Function space

Vector Space videos
https://www.youtube.com/watch?v=ozwodzD5bJM

Basis set for a given Vector space

Minimal set of independent vectors whose linear combinations can produce all the vectors in the vector space

For R^2 we need 2 independent 2-tuple vectors

For R^3 we need 3 independent 3-tuple vectors

For R^n we need n independent n-tuple vectors

Subspace of Vector space definition

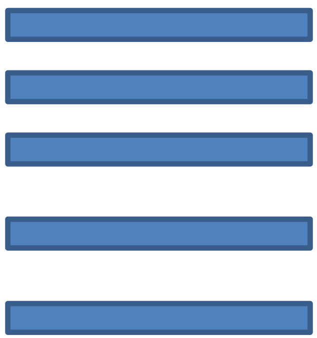
Layman definition:
A subset of vectors of a full vector space with all the property of a vector space

Set of all points of any line (note : line is a set of point vectors) passing through origin in R^2 is a subspace of R^2

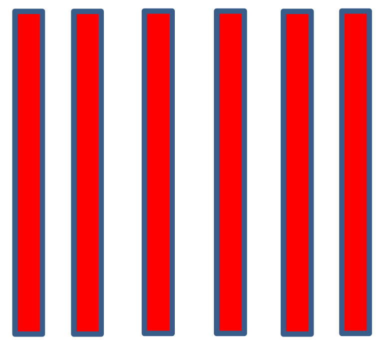
Set of points on any Plane or a line (note : plane is a set of point vectors) passing through origin in R^3 is a subspace of R^3

Central Concept in Linear Algebra - Four Fundamental Subspaces

Vector Spaces associated with a Matrix



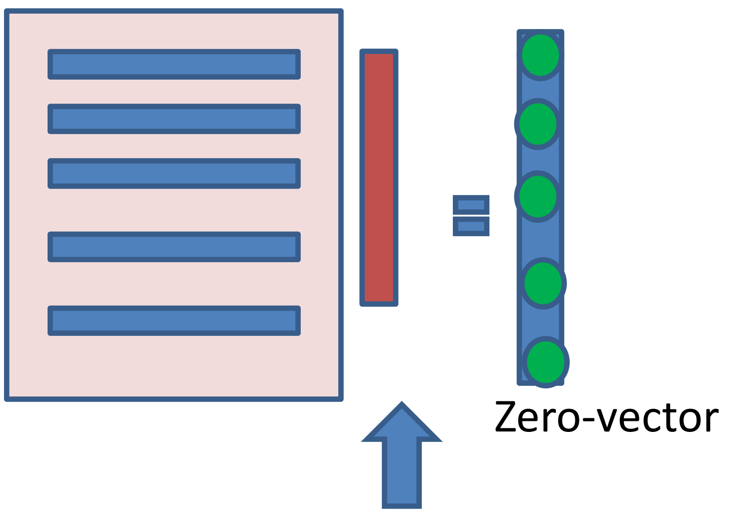
Row-space
=vector space generated by
All possible LC of row-vectors



Column-space

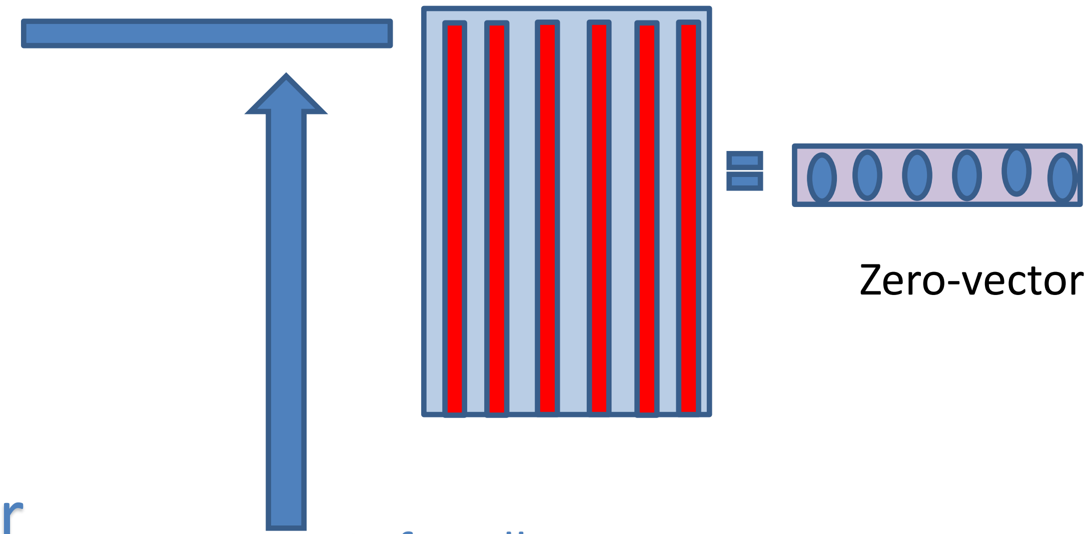
Vector Space generated by
All possible LC of column vectors

Central Concept in Linear Algebra - Four Fundamental Subspaces



Right-nullspace vector

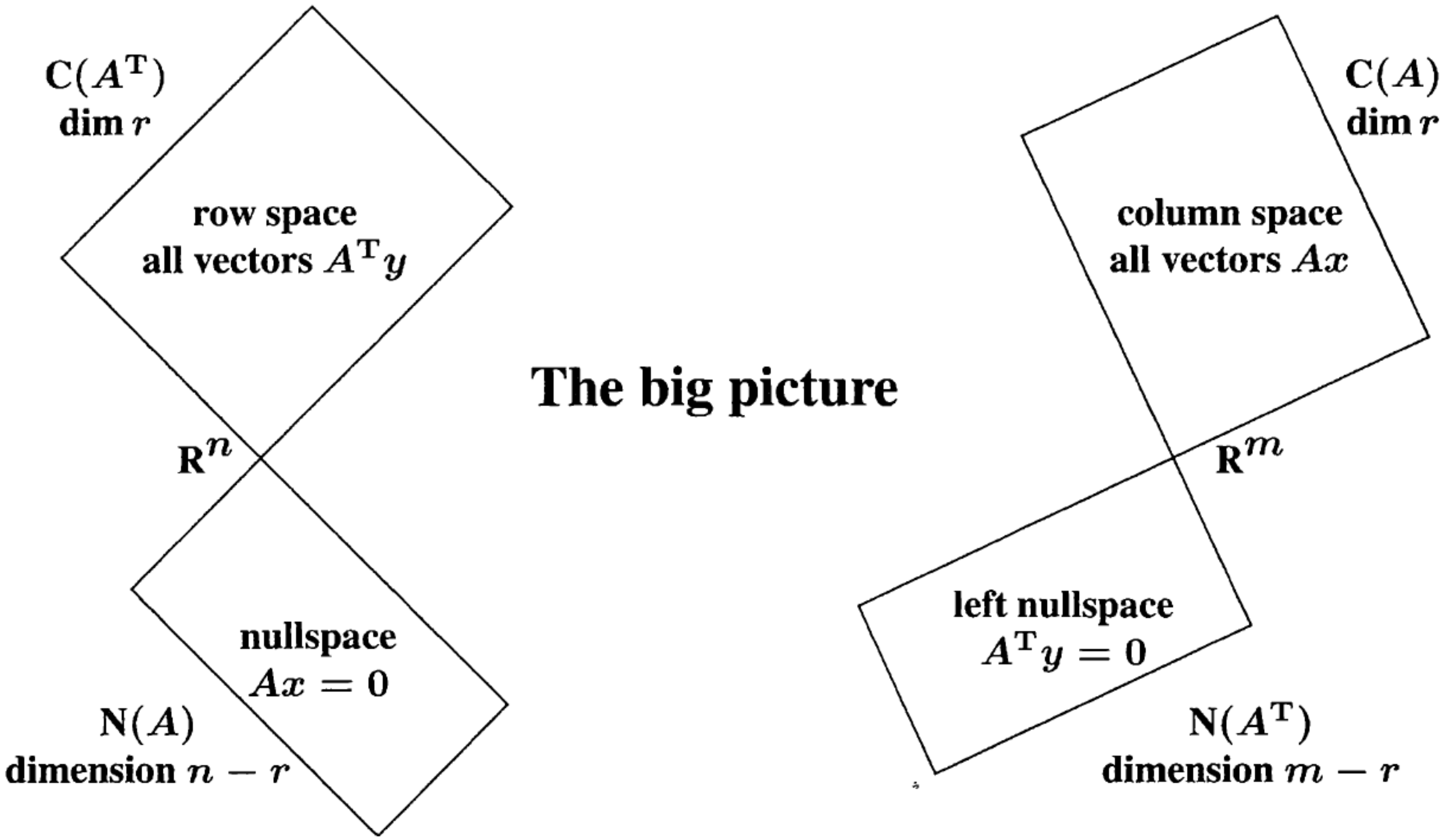
Vector-space generated by LC of all such vectors



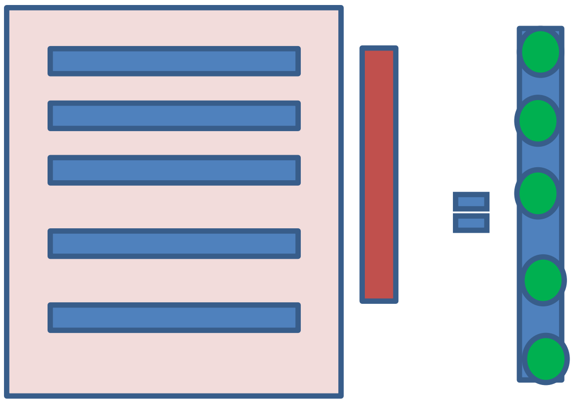
Left-nullspace vector

Vector-Space generated by LC of all such vectors

Central Concept in Linear Algebra
- Four Fundamental Subspaces



Fundamental Theorem of Linear Algebra
- Relation between the fundamental subspaces



$$\begin{bmatrix} \bar{r}_1 \\ \bar{r}_2 \\ \vdots \\ . \\ \bar{r}_n \end{bmatrix} \mathbf{X} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ . \\ 0 \end{bmatrix}$$

From AX=0, we get:

Each row of A is orthogonal to X

From $A^T Y=0$, we get:

Each row of A^T is orthogonal to Y

i.e., Each column of A is orthogonal to Y

Fundamental Theorem of Linear Algebra
- Relation between the fundamental subspaces

Each row of A is orthogonal to X

$\rightarrow$ Linear combinations of row vectors of A are orthogonal to X

$\rightarrow$ Row Space of A is orthogonal to Null space of A

Each column of A is orthogonal to Y

$\rightarrow$ Linear combinations of column vectors of A are orthogonal to Y

$\rightarrow$ Column Space of A is orthogonal to Left Null space of A or Null space of A^T

Eigenvalues and Eigenvectors

Is there any special vector x (that specify a direction)

Such that :

$$Ax = \lambda x$$

Is there any special direction for A such that it does not do rotation but do only scaling?.

Answer is Yes

And such a direction is called an eigenvector

And the factor by which there is the scaling is called the eigenvalue

Eigenvalues and eigenvectors

- Let $\mathbf{A}$ be a $n \times n$ matrix. The vector $\mathbf{v}$ that satisfies

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

for some scalar λ is called the **eigenvector** of $\mathbf{A}$ and λ is the **eigenvalue** corresponding to the eigenvector $\mathbf{v}$.

Application

Let A be a 2x2 matrix

Let X be a set of points on a unit circle and let $x \in X$

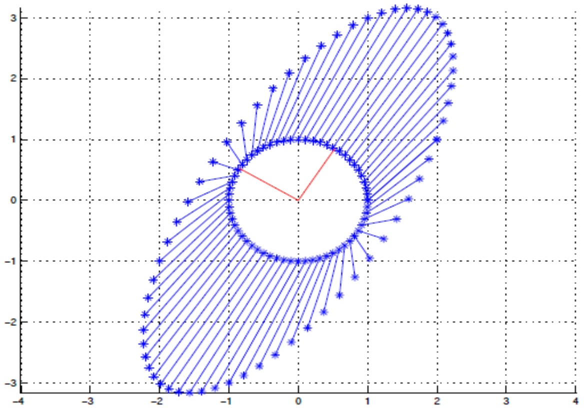
Let Y be a set of points such that $y = Ax$, $y \in Y$

In general y vector is a scaled and rotated version of x

But we note that there are special x's (directions) such that $Ax = \lambda x$

For 2×2 matrix, there are two such directions(in general, except when $\lambda_1 = \lambda_2$)

Example: If $A = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$



MATLAB Command

- eig(A)   (returns eigenvalues)

- [evec   eval] = eig(A)

  (returns eigenvalues and eigenvectors)

- Matlab provide eigen vectors with unit norm

Properties of Eigenvalues

1. **A square matrix of order n will have n eigenvalues. The eigenvalues can be real or complex conjugates**

This is because roots of an nth degree polynomial equation has n roots and also these roots are either real or if complex appears in pairs as complex conjugates.

Properties of Eigenvalues

2. Sum of Eigenvalues of a matrix = Trace of the matrix (sum of diagonal elements of matrix)

3. Product of Eigenvalues of a matrix = Determinant of the matrix

4. If λ_1 and λ_2 are the eigenvalues of a square matrix A of order 2, then:

(a) ${\lambda_1}^2$ and ${\lambda_2}^2$ will be the eigenvalues of the matrix A^2
(b) ${\lambda_1}^3$ and ${\lambda_2}^3$ will be the eigenvalues of the matrix A^3
(c) ${\lambda_1}^n$ and ${\lambda_2}^n$ will be the eigenvalues of the matrix A^n

But the unit magnitude eigenvectors will be same for all these matrices

Properties of Eigenvalues

5. If λ_1 and λ_2 are the eigenvalues of a square matrix A of order 2, then, $k\lambda_1$ and $k\lambda_2$ are the eigenvalues of the matrix kA.
6. If λ_1 and λ_2 are the eigenvalues of a non-singular matrix A of order 2, then, $\frac{1}{\lambda_1}$and $\frac{1}{\lambda_2}$ are the eigenvalues of the matrix A^{-1}.

   (But the unit magnitude eigenvectors will be same of all these matrices)
7. Eigenvalues of A and A^T are same
8. Eigenvalues of a symmetric matrix is real and their eigenvectors are orthogonal.
9. Eigenvalues of a skew symmetric matrix is purely imaginary or zero.
10. Eigenvalues of an orthogonal matrix has magnitude 1.
11. If all row sums or column sums are same for a matrix A, then that sum is an eigenvalue of A.

Properties of Eigenvalues

12. If eigenvalues are distinct, the corresponding eigenvectors will be linearly independent.
    Hence the eigenvectors of an $n \times n$ matrix with distinct eigenvalues will form a basis for the Euclidean Space R^n.
    (If eigenvalues are repeated then eigenvectors may or may not form a basis of eigenvectors for R^n.)
13. Similar matrices have same eigenvalues.
    (Two matrices A and B are similar, if their exist an invertible matrix P such that $B=P^{-1}AP$ or there is an invertible matrix M such that $A=M^{-1}BM$).
14. AB and BA are similar and so have same non-zero eigenvalues.
15. A^TA and AA^T have same non-zero eigenvalues.
16. Cayley-Hamilton Theorem: Every matrix satisfies its own characteristic equation.

Properties of Eigenvalues

17. **Diagonalization:**
    $\Lambda = X^{-1}AX$ or $A = X\Lambda X^{-1}$
    where X is the matrix with column vectors as eigenvectors of A and Λ is the diagonal matrix with eigenvalues of A as diagonals.
18. For symmetric matrices, eigenvalues are real and eigenvectors are orthogonal, hence $A = X\Lambda X^{-1}$ will reduce to $S = Q\Lambda Q^T$, where Q is an orthogonal matrix. This is called **Spectral Decomposition of S.**
19. Eigenvalues of a positive definite matrix is all positive and the eigenvectors of a positive definite matrix is orthogonal.
20. A^TA and AA^T will have either positive eigenvalues or zero eigenvalues.
    (If all columns of a matrix A are independent, then A^TA will be a positive definite matrix. If some columns of A are dependent, then A^TA will be a semi-positive definite matrix).

Matrix decompositions

- There are numerous examples of useful matrix decompositions:
- Matrix **factorization** is the same thing as matrix decomposition (e.g. NMF = nonnegative matix factorization, $\mathbf{V} = \mathbf{WH}$, all elements nonnegative.)

Important Decompositions that we learned

- CR Decomposition
- LU Decomposition
- QR Decomposition
- Spectral Decomposition
- **SVD**
- QS Decomposition

Singular Value Decomposition

Spectral Decomposition for square symmetric matrices

$$\underset{n\times n}{S} = \underset{n\times n}{Q}\,\underset{n\times n}{\Lambda}\,\underset{n\times n}{Q^T} = \lambda_1 q_1 q_1^T + \lambda_2 q_2 q_2^T + \cdots + \lambda_n q_n q_n^T$$

$$\boxed{\text{Square matrices:- } \quad A = X D X^{-1}}$$

$$\Sigma \rightarrow \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Singular Value Decomposition for rectangular matrices

$$A = U \Sigma V^T = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T$$

Orthogonal Diagonal Orthogonal.

$$A_{m\times n} = \underset{m\times m}{U}\,\underset{m\times n}{\Sigma}\,\underset{n\times n}{V^T}$$

$$\bar{y} = A\bar{x}$$

$$= \underbrace{Q \underbrace{\Lambda \underbrace{Q^T \bar{x}}_{\text{Rotation}}}_{\text{Scaling}}}_{\text{Rotation.}}$$

Singular Value Decomposition

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}V^T_{n\times n}$$

If $\Sigma =$

$$\Sigma = \begin{bmatrix}\sigma_1 & 0\\ 0 & \sigma_2\\ 0 & 0\end{bmatrix} \text{ or } \begin{bmatrix}\sigma_1 & 0 & 0 & 0\\ 0 & \sigma_2 & 0 & 0\end{bmatrix}$$

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T$$

$$\underset{2\times 3}{\Sigma^T}\underset{3\times 2}{\Sigma} = \begin{bmatrix}\sigma_1^2 & 0\\ 0 & \sigma_2^2\end{bmatrix}$$

$$\Sigma^T\Sigma = \underset{4\times 4}{\begin{bmatrix}\sigma_1^2 & 0 & 0 & 0\\ 0 & \sigma_2^2 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}}$$

$$\underbrace{A^TA}_{\text{Symm. matrix}} = (U\Sigma V^T)^T(U\Sigma V^T)$$

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

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

$$= V\Sigma^T I\Sigma V^T \quad (\because U^TU = I, \text{ as } U \text{ is orthogonal})$$

$$= V\underbrace{\Sigma^T\Sigma}_{\text{Diagonal matrix.}} V^T = VDV^T \longrightarrow \text{similar to spectral dec. } S = Q\Lambda Q^T$$

Post multiply by V,

$$A^TAV = V\Sigma^T\Sigma \Rightarrow \text{similar to } AX = X\Lambda$$

$\Rightarrow$ V contains eigenvectors of A^TA in its columns

$\Sigma^T\Sigma$ contains eigenvalues of A^TA in its diagonal.

$\Rightarrow \Sigma$ has root of λs of A^TA in its diagonal (real nos)

Singular Value Decomposition

$$AA^T = (U\Sigma V^T)(U\Sigma V^T)^T$$
$$= U\Sigma \underbrace{V^T V}\Sigma^T U^T$$
$$= U\Sigma\Sigma^T U^T \quad (\text{Spectral Decomposition of } AA^T)$$
$$= UDU^T$$

$\Rightarrow U \rightarrow$ eigenvectors of AA^T

$\Sigma\Sigma^T \rightarrow$ eigenvalues of AA^T

Non-zero λs of A^TA & AA^T are same.

A^TA is +ve definite $\Rightarrow$ +ve λs $(\sigma_1^2, \sigma_2^2, \ldots \geq 0)$

$\Rightarrow$ real $\sigma_1, \sigma_2, \ldots$

$\sigma_1, \sigma_2, \ldots$ are called singular values of A.

$\sigma_1^2, \sigma_2^2, \ldots$ are eigenvalues of A^TA

Singular Value Decomposition

$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$

$\Sigma^T \Sigma \mid \Sigma \Sigma^T$

$\Sigma^T \Sigma$

$$A_{m\times n} = U_{m\times m} \Sigma_{m\times n} V^T_{n\times n}$$

$U \to$ Orthogonal matrix with unit eigenvectors of AA^T in columns.

$u_1, u_2, \ldots u_n \to$ column vectors of U

$V \to$ Orthogonal matrix with unit eigenvectors of A^TA in columns.

$v_1^T, v_2^T, \ldots v_n^T \to$ row vectors of V^T

$\Sigma_{m\times n} \to$ leading diagonal elts. as square roots of eigenvalues of A^TA & AA^T

$$\boxed{\begin{aligned} Av_1 &= \sigma_1 u_1 \\ Av_2 &= \sigma_2 u_2 \\ &\vdots \\ Av_n &= \sigma_n u_n \end{aligned}}$$

$$AV = U\Sigma$$
$$A = U\Sigma V^T$$

Singular Value Decomposition

① $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}_{2\times 3}$. Find SVD for A.

$$A^TA = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} \rightarrow 3\times 3 \text{ symmetric}$$

$\longrightarrow \lambda_s = 3, 1, 0$

$\bar{x}_\lambda = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$

$\bar{x}_u = \underbrace{\begin{pmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}}_{V_1}, \underbrace{\begin{pmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{pmatrix}}_{V_2}, \underbrace{\begin{pmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}}_{V_3}$

$$AA^T = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}_{\substack{2\times 2 \\ \text{symmetric}}}$$

$\longrightarrow \lambda_s = 3, 1$

$\bar{x}_\lambda = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$\bar{x}_u = \underbrace{\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}}_{u_1}, \underbrace{\begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}}_{u_2}$

$$A = U \Sigma_{2\times 3} V^T$$

$$= \begin{array}{c} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \\ \lambda=3 \quad \lambda=1 \end{array} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{array}{c} \begin{bmatrix} 1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix} \end{array} \begin{array}{l} \lambda=3 \\ \lambda=1 \\ \lambda=0 \end{array}$$

$\sigma_1 = \sqrt{3}$

$\sigma_2 = 1$

Singular Value Decomposition

$\sigma_1 = \sqrt{\lambda \text{ of } A^TA / AA^T}$

$\sigma_2 = \text{"}$

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 0 & 0 \end{bmatrix}_{3\times2} = U_{3\times3}\,\Sigma_{3\times2}\,V^T_{2\times2}$$

$$= \begin{bmatrix} | & | & | \\ u_1 & u_2 & u_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} - & v_1 & - \\ - & v_2 & - \end{bmatrix}$$

$$A^TA = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}_{2\times2} \longrightarrow \lambda s \text{ of } A^TA \cdot \{\underset{v_1}{6.8541}, \underset{v_2}{0.1459}\}$$

$\sigma_1 = \sqrt{6.8541}$

$\sigma_2 = \sqrt{0.1459}$

$$AA^T = \begin{bmatrix} 2 & 3 & 0 \\ 3 & 5 & 0 \\ 0 & 0 & 0 \end{bmatrix}_{3\times3} \rightarrow \lambda s \text{ of } AA^T = \{6.8541, 0.1459, 0\}$$

Singular Value Decomposition

$$A = U\Sigma V^{\mathrm{T}} = \begin{pmatrix} \text{orthogonal} \\ U \text{ is } m \times m \end{pmatrix} \begin{pmatrix} m \times n \text{ singular value matrix} \\ \sigma_1, \ldots, \sigma_r \text{ on its diagonal} \end{pmatrix} \begin{pmatrix} \text{orthogonal} \\ V \text{ is } n \times n \end{pmatrix}.$$

Requirements: None. This *Singular Value Decomposition* (SVD) has the eigenvectors of AA^{T} in U and eigenvectors of $A^{\mathrm{T}}A$ in V; $\sigma_i = \sqrt{\lambda_i(A^{\mathrm{T}}A)} = \sqrt{\lambda_i(AA^{\mathrm{T}})}$.

Those singular values are $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$. By column-row multiplication

$$A = U\Sigma V^{\mathrm{T}} = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^{\mathrm{T}} + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^{\mathrm{T}}.$$

If S is symmetric positive definite then $U = V = Q$ and $\Sigma = \Lambda$ and $S = Q\Lambda Q^{\mathrm{T}}$.

Symmetric S	$S = Q\Lambda Q^{\mathrm{T}}$	$= \lambda_1 \boldsymbol{q}_1 \boldsymbol{q}_1^{\mathrm{T}} + \lambda_2 \boldsymbol{q}_2 \boldsymbol{q}_2^{\mathrm{T}} + \cdots + \lambda_r \boldsymbol{q}_r \boldsymbol{q}_r^{\mathrm{T}}$
Any matrix A	$A = U\Sigma V^{\mathrm{T}}$	$= \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^{\mathrm{T}} + \sigma_2 \boldsymbol{u}_2 \boldsymbol{v}_2^{\mathrm{T}} + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^{\mathrm{T}}$

Singular Value Decomposition

MATLAB command for Singular Value Decomposition

$A = UZV'$

```
>[U,Z,V] = svd(A)
```

% produces a diagonal matrix Z, of the same dimension as A and with nonnegative diagonal elements in decreasing order, and unitary matrices U and V so that A = U*Z*V'.

```
>Z = svd(A)
```

% returns a vector containing the singular values.

AMRITA VISHWA VIDYAPEETHAM UNIVERSITY Established u/s 3 of UGC Act 1956

Singular Value Decomposition

Frobenius norm of $A=[a_{ij}]$

$= \sqrt{\sum\sum a_{ij}^2}$

(a) Why is the trace of $A^T A$ equal to the sum of all a_{ij}^2 ?

$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$; $A^T A = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{11}^2 + a_{21}^2 & a_{11}a_{12} + a_{21}a_{22} \\ a_{12}a_{11} + a_{22}a_{21} & a_{12}^2 + a_{22}^2 \end{bmatrix}$

$\boxed{\text{Trace}(A^T A) = a_{11}^2 + a_{21}^2 + a_{12}^2 + a_{22}^2} = (\text{Frobenius norm})^2$

** $\text{Trace}(A^T A)$ = Sum of all eigenvalues of $A^T A$.

$= \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2$.

** For $A_{m \times n}$, $\boxed{\sigma_1^2 + \sigma_2^2 + \cdots = (\text{Frobenius norm})^2}$

(b) For every rank-one matrix, why is $\sigma_1^2 = \text{sum of all } a_{ij}^2$?

$A \rightarrow$ rank 1 matrix.

A will have only one singular value,

$\sigma_1^2 = \sum_i \sum_j a_{ij}^2$.

Amrita Vishwa Vidyapeetham University Established u/s 3 of UGC Act 1956

Singular Value Decomposition

Qn.) Find the sum of squares of all singular values of the following matrices:- without finding the singular values.

(a) $A = \begin{bmatrix} 1 & 6 & 7 & 5 \\ 3 & 2 & 5 & 1 \end{bmatrix}$; (b) $B = \begin{bmatrix} 3 & 5 & 6 \\ 1 & 2 & 1 \\ 2 & 1 & -1 \\ 0 & 3 & 0 \end{bmatrix}$

Soln:-

(a) A will have 2 singular values

$$\sigma_1^2 + \sigma_2^2 = 1^2 + 6^2 + 7^2 + 5^2 + 3^2 + 2^2 + 5^2 + 1^2 = 150$$

```
>> A=[1,6,7,5;3,2,5,1];
>> FN=norm(A,'fro')
FN =
  12.2474
>> FN^2
ans =
  150
```

(b) B will have 3 singular values.

$$\sigma_1^2 + \sigma_2^2 + \sigma_3^2 = (\text{Frob. norm})^2$$
$$= 91$$

SVD to find an orthogonal basis for the four fundamental subspaces of a matrix

Given the SVD of an m×n matrix with rank r as $A=U\Sigma V^T$, with column vectors of U as $\boldsymbol{u_1}, \boldsymbol{u_2},\ldots, \boldsymbol{u_m}$ and column vectors of V as $\boldsymbol{v_1}, \boldsymbol{v_2},\ldots,\boldsymbol{v_n}$.

The orthonormal basis for the four fundamental subspaces of A are as follows:

$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$	is an orthonormal basis for the **column space**
$\boldsymbol{u}_{r+1}, \ldots, \boldsymbol{u}_m$	is an orthonormal basis for the **left nullspace** $\boldsymbol{N}(A^{\mathrm{T}})$
$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$	is an orthonormal basis for the **row space**
$\boldsymbol{v}_{r+1}, \ldots, \boldsymbol{v}_n$	is an orthonormal basis for the **nullspace** $\boldsymbol{N}(A)$.

Singular Value Decomposition

MATLAB command for Singular Value Decomposition

```
>[U,Z,V] = svd(A)
```

% produces a diagonal matrix Z, of the same dimension as A and with nonnegative diagonal elements in decreasing order, and unitary matrices U and V so that A = U*Z*V'.

```
>Z = svd(A)
```

% returns a vector containing the singular values.

Singular Value Decomposition

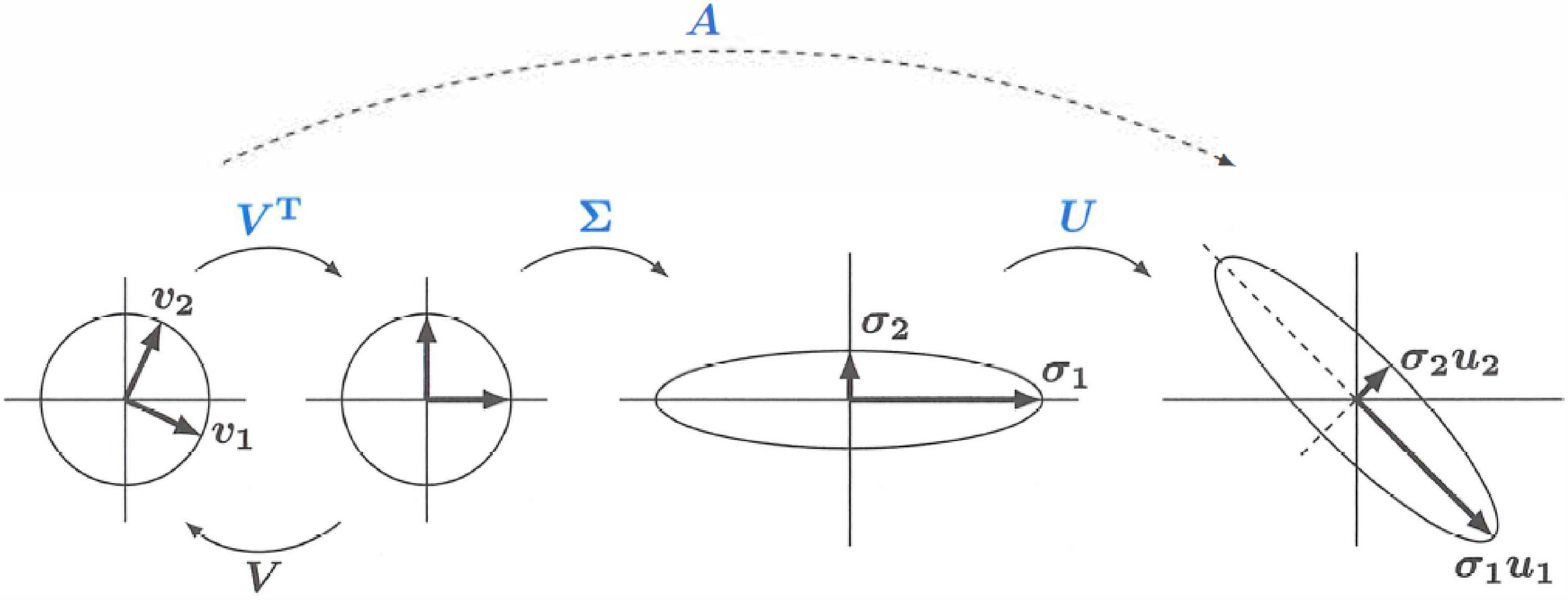
Using MATLAB find the singular values of A (not using the command svd(A)), and the singular vectors u1 and u2.

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 2.5 \\ 3 & 4 \\ 4 & 4.2 \\ -1 & -1.5 \\ -2 & -2.8 \\ -3 & -3.5 \\ -4 & -3.9 \end{bmatrix}$$

```
A=[1 1;2 2.5;3 4;4 4.2;-1 -1.5;-2 -2.8;-3 -3.5;4 -3.9];
[v, lamda]=eig(A'*A);
v1=v(:,1);
v2=v(:,2);
lamdas=diag(lamda);
lamda1=lamdas(1);
lamda2=lamdas(2);
Total_variation=lamda1+lamda2;
Frobnormsquare =(norm(A,'fro'))^2;
sigma1=norm(A*v1); %sigma1 is also same as root of lamda1
sigma2=norm(A*v2); %sigma2 is also same as root of lamda2
u1=A*v1/sigma1;
u2=A*v2/sigma2;
```

Geometry of SVD

The SVD separates a matrix into three steps: (**orthogonal**)$\times$(**diagonal**)$\times$(**orthogonal**). Ordinary words can express the geometry behind it: (**rotation**)$\times$(**stretching**)$\times$(**rotation**). $U\Sigma V^{\mathrm{T}}\boldsymbol{x}$ starts with the rotation to $V^{\mathrm{T}}\boldsymbol{x}$. Then Σ stretches that vector to $\Sigma V^{\mathrm{T}}\boldsymbol{x}$, and U rotates to $A\boldsymbol{x} = U\Sigma V^{\mathrm{T}}\boldsymbol{x}$. Here is the picture.



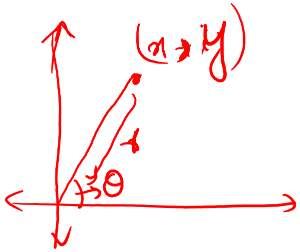
Important Decompositions

- **Polar** Decomposition / QS Decomposition

stretch, rotation

$$z = x + iy = \underline{r}\, e^{i\theta} = r(\cos\theta + i\sin\theta)$$

$$= r\cos\theta + i\sin\theta$$



$$\text{Matrix} = \underset{\text{rot.}}{Q}\ \underset{\text{stretch}}{S}$$

$$A = QS = (\text{orthogonal matrix } Q)\ (\text{symmetric positive definite matrix } S).$$

Every real square matrix can be factored into $A = QS$, where Q is ***orthogonal*** and S is ***symmetric positive semidefinite***. If A is invertible, S is positive definite.

Derivation of Polar Decomposition from SVD

$$A = U\Sigma V^T = U I \Sigma V^T$$
$$= U(V^TV)\Sigma V^T$$
$$= \underbrace{(UV^T)}_{\text{Orthogonal}}\ \underbrace{(V\Sigma V^T)}_{\text{Symmetric \& +ve definite}}$$
$$= Q \times S$$

V, U orthogonal $\Rightarrow U, V^T$ orthogonal $\Rightarrow UV^T$ orthogonal

(transpose of orth. matrix is orth.) (product of 2 orth. matrices is orthogonal)

V is orthogonal, Σ diagonal, $S = V\Sigma V^T$ ($S = Q^T\Sigma Q$)

S is symmetric & also +ve def

Also,
$$A = U\Sigma V^T$$
$$= U\Sigma(U^TU)V^T$$
$$= \underbrace{(U\Sigma U^T)}_{M}\ \underbrace{(UV^T)}_{Q}$$

M = Sym. +ve definite; Q orthogonal.

$A = MQ$

where $M = U\Sigma U^T$

& $Q = UV^T$

is also a decomposition with +ve semi def. matrix first & then the orthogonal matrix.

Importance of Polar Decomposition:

- $Q=UV^T$ is orthogonal as it is product of orthogonal matrices U and V^T.

- $S=V\sum V^T$ is positive semidefinite (positive definite if A is invertible), since V is orthogonal and $\sum$ has positive diagonal elements.
  ($S=V\sum V^T$ would be the spectral decomposition)

- Polar Decomposition separates the **rotation** (in Q) from **stretching** (in S)

1. Find the polar decomposition for the matrix: $A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}$.

Find SVD of A :- $A = U \Sigma V^T$, where $U = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}$; $\Sigma = \begin{bmatrix} \sqrt{45} & 0 \\ 0 & \sqrt{5} \end{bmatrix}$; $V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

Polar Decomposition of A is $A = QS$

where $Q = UV^T = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix} \times \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \frac{1}{\sqrt{20}} \begin{bmatrix} 4 & -2 \\ 2 & 4 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}$

$$S = V \Sigma V^T = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{45} & 0 \\ 0 & \sqrt{5} \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} \sqrt{45} & -\sqrt{5} \\ \sqrt{45} & \sqrt{5} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \sqrt{45}+\sqrt{5} & \sqrt{45}-\sqrt{5} \\ \sqrt{45}-\sqrt{5} & \sqrt{45}+\sqrt{5} \end{bmatrix}$$

$$= \frac{\sqrt{5}}{2} \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix} = \sqrt{5} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Verify, $QS = A$

1. Find the polar decomposition for the matrix: $A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}$ using MATLAB.

```
>> A=[3,0;4,5];
>> [U,Z,V]=svd(A);
>> Q=U*V';
>> S=V*Z*V'
>> Q*S  %Verify if it is A
```

Given a data, represented by a matrix A, In which direction the variation is maximum?

Or, In which direction the data is maximally spread?.

Let $\boldsymbol{v} \in \boldsymbol{R}^2$ be the unit vector in the direction in which the variation is maximum.

$\mathrm{T}_v = (A\boldsymbol{v})^T (A\boldsymbol{v})$ represent total variation(or measure of spread) along v

How can we formulate the problem as an optimization problem?

$$\underset{v}{\text{Maximize}} \; v^T A^T A v$$
$$subject \text{ to } v^{\mathrm{T}} v = 1$$

Lagrangian function

$$L(v,\lambda) = v^T A^T A v - \lambda\left(v^T v - 1\right)$$

$$L(v,\lambda) = v^T A^T A v - \lambda\left(v^T v - 1\right)$$

Necessary condition for optimality

$$\frac{\partial L}{\partial v} = 0,$$

i.e., $2A^T A v - 2\lambda v = 0 \Rightarrow A^T A v = \lambda v$

The necessary condition implies that along the direction v, i.e., along the direction of eigenvector there is optimum variation

Substituting the necessary condition in the objective, we get the value as λ, which will be two either $\lambda_1\ or \lambda_2$. Since the objective is maximize the largest of $\lambda_1\ and\ \lambda_2$ is the optimum

$$v^T A^T A v = v^T \lambda v = \lambda v^T v = \lambda$$

Hence the eigenvector v corresponding to the largest eigenvalue λ of the matrix $A^T A$ gives the direction along which there is maximum variation.

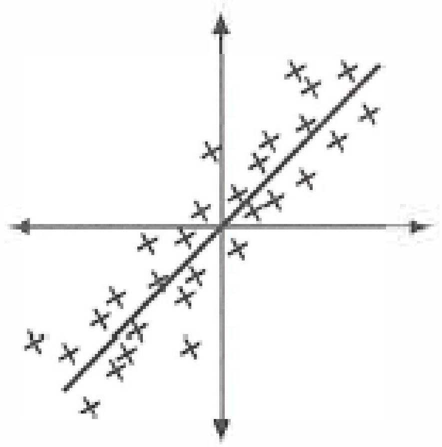
Principal Component Analysis

1 Data often comes in a matrix : n samples and m measurements per sample.

2 Center each row of the matrix A by subtracting the mean from each measurement.

3 The SVD finds combinations of the data that contain the most information.

4 Largest singular value $\boldsymbol{\sigma_1}$ $\leftrightarrow$ greatest variance $\leftrightarrow$ most information in $\boldsymbol{u_1}$.



A is $\mathbf{2} \times \boldsymbol{n}$ (large nullspace)

AA^{T} is $\mathbf{2} \times \mathbf{2}$ (small matrix)

$A^{\mathrm{T}}A$ is $\boldsymbol{n} \times \boldsymbol{n}$ (large matrix)

Two singular values $\boldsymbol{\sigma_1} > \boldsymbol{\sigma_2} > \mathbf{0}$

Principal Component Analysis

The Essentials of Principal Component Analysis (PCA)

PCA gives a way to understand a data plot in dimension m = the number of measured variables (here age and height). Subtract average age and height ($m = 2$ for n samples) to center the m by n data matrix A. *The crucial connection to linear algebra* is in the singular values and singular vectors of A. Those come from the eigenvalues $\lambda = \sigma^2$ and the eigenvectors $\boldsymbol{u}$ of the sample covariance matrix $S = AA^{\mathrm{T}}/(n-1)$.

- The total variance in the data is the sum of all eigenvalues and of sample variances s^2:
  Total variance $\boldsymbol{T} = \boldsymbol{\sigma_1^2} + \cdots + \boldsymbol{\sigma_m^2} = \boldsymbol{s_1^2} + \cdots + \boldsymbol{s_m^2} =$ **trace** (*diagonal sum*).
- The first eigenvector $\boldsymbol{u}_1$ of S points in the most significant direction of the data. That direction accounts for (or *explains*) a fraction σ_1^2/T of the total variance.
- The next eigenvector $\boldsymbol{u}_2$ (orthogonal to $\boldsymbol{u}_1$) accounts for a smaller fraction σ_2^2/T.
- Stop when those fractions are small. You have the R directions that explain most of the data. The n data points are very near an R-dimensional subspace with basis $\boldsymbol{u}_1$ to $\boldsymbol{u}_R$. These $\boldsymbol{u}$'s are the **principal components** in m-dimensional space.
- R is the "effective rank" of A. The true rank r is probably m or n: full rank matrix.

Class Test

In all the Questions, Y stands for Last Two Digits of your Registration Number
Replace Y where ever you see it

Write Question Number and Answer

1

Find Eigenvalues of following matrix

$$A=\begin{pmatrix} 1 & 0 & 0 \\ 0 & Y & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Your answer

Eigenvalues are ------------, ------------------. -----------------

2

Find Eigenvalues of following matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -Y & 4 \\ 0 & 0 & Y \end{pmatrix}$$

Your answer

Eigenvalues are -----------, ----------------. ----------------

3

Find Eigenvalues of following matrix

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 3 & Y & 0 \\ 4 & 5 & -Y \end{pmatrix}$$

Your answer

Eigenvalues are ------------, ------------------. ------------------

4

Find Eigenvalues of following matrix

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & Y+1 & 0 \\ 2 & 0 & 0 \end{pmatrix}$$

Your answer

Eigenvalues are -----------, ----------------. ----------------

Ans :

5

Find Eigenvalues of following matrix

$$A=\begin{pmatrix} 0 & 0 & Y+1 \\ 0 & Y+1 & 0 \\ Y+1 & 0 & 0 \end{pmatrix}$$

Your answer

Eigenvalues are -----------, ----------------. ----------------

6

Find Eigenvalues of following matrix

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & Y+1 & 0 \\ 0 & Y+1 & 0 & 0 \\ 4 & 0 & 0 & 0 \end{pmatrix}$$

Your answer

Eigenvalues are --------, ------------. -----------,-----------

7

Find Eigenvalues of following matrix

$$A = \begin{matrix} 0 & 0 & 0 & Y+1 \\ 0 & 0 & Y+1 & 0 \\ 0 & Y+1 & 0 & 0 \\ Y+1 & 0 & 0 & 0 \end{matrix}$$

Your answer

Eigenvalues are --------, ------------. ----------,----------

8

Find Determinant of the following matrix

$$A = \begin{pmatrix} 1 & -(Y+1) & 0 \\ 1 & Y+1 & 0 \\ 3 & 5 & 5 \end{pmatrix}$$

Your answer

Determinant is ------------

9

Find Eigenvalues of the following matrix

$$A=\underbrace{\begin{pmatrix} 1 & -(Y+1) & 0 \\ 1 & Y+1 & 0 \\ 3 & 5 & 5 \end{pmatrix}}_{Look\ \text{at coloumnwise sum}}$$

Use trace and determinant

$$\lambda_2+\lambda_3=a-----(1)$$
$$\lambda_2\lambda_3=b$$
$$(\lambda_2-\lambda_3)^2=(\lambda_2+\lambda_3)^2-4\lambda_2\lambda_3=a^2-4b$$
$$(\lambda_2-\lambda_3)=\sqrt{a^2-4b}----(2)$$
$$(1)+(2)\text{ gives}$$
$$2\lambda_2=a+\sqrt{a^2-4b}$$
$$\lambda_2=\frac{a+\sqrt{a^2-4b}}{2}$$
$$\lambda_3=\frac{a-\sqrt{a^2-4b}}{2}$$

Your answer

Eigenvalues are -----------, -----------, -----------,

10

Find Eigenvalues of the following matrix

$$A=\begin{pmatrix} Y+1 & Y+1 & Y+1 \\ Y+1 & Y+1 & Y+1 \\ Y+1 & Y+1 & Y+1 \end{pmatrix}$$

Your answer

Eigenvalues are ------------, ------------, ------------,

11

Find Eigenvector corresponding to a non-zero eigenvalue

$$A = \begin{pmatrix} Y+1 & Y+1 & Y+1 \\ Y+1 & Y+1 & Y+1 \\ Y+1 & Y+1 & Y+1 \end{pmatrix}$$

Your answer

Eigenvector is $(\text{-----}, \text{-------}, \text{---------})^T$

12

Find Eigenvector corresponding to λ=5

$$A = \begin{pmatrix} 5 & 2 & -2 \\ 5 & -Y & Y \\ 3 & 2 & 0 \end{pmatrix}$$

Your answer

Eigenvector is (-----, -------, ---------)^T

13

Find Eigenvalues *of* following magic square

$$A = \begin{pmatrix} 7 & 0 & 5 \\ 2 & 4 & 6 \\ 3 & 8 & 1 \end{pmatrix}; \ \text{det(A)}= -288; \ \text{Trace(A)=12}; \quad \text{-}288/12 = -24$$

Your answer

Eigenvalues are $------,--------.------$

14

Eigenvalues of a 3×3 matrix A are $Y+1, Y+2, Y+3$

What is Trace and determinant of that matrix

Your answer

Trace = -------, Determinant = --------

15

Eigenvalues *of* a 3×3 matrix A are Y,Y+1,Y+2

What is Trace and determinant of the matrix A^2

Your answer

Trace = $-------$, Determinant = $--------$

16

Find a vector orthogonal to rowvectors of following Matrix A

$$A = \begin{pmatrix} 1 & Y & 1 \\ 2 & Y & 0 \end{pmatrix}$$

Your Answer

A *vector* orthogonal to rowvectors of $\begin{pmatrix} 1 & - & 1 \\ 2 & - & 0 \end{pmatrix}$ is (-------,-----,-----)^T

17

$Find$ Eigenvector $corresponding$ to eigenvalue 12

$$A = \begin{pmatrix} 7 & 0 & 5 \\ 2 & 4 & 6 \\ 3 & 8 & 1 \end{pmatrix};$$

$Your$ answer

Eigenvector is $(------,--------,------)^T$

18

Find Charactetistic Polynomial of Matrix A in variable λ

$$A = \begin{pmatrix} 1 & -2 & 3 \\ -1 & 3 & -2 \\ -2 & 5 & -4 \end{pmatrix}.$$

Your answer is :

$$A = \begin{pmatrix} x & a & c \\ b & y & e \\ d & f & z \end{pmatrix} \Rightarrow (\lambda I - A) = \begin{pmatrix} \lambda - x & -a & -c \\ -b & \lambda - y & -e \\ -d & -f & \lambda - z \end{pmatrix}$$

$$|(\lambda I - A)| = \lambda^3 - \underbrace{(x + y + z)}_{Tr(A)} \lambda^2 +$$

$$\left(\underbrace{(xy + yz + zx)}_{diagonal} - \underbrace{(ab + cd + ef)}_{off\_diagonal} \right) \lambda - \det(A)$$

19

$$Eigenvalues \text{ of } A=\begin{pmatrix} 3 & -3 & 4 \\ 2 & -4 & 8 \\ 2 & -4 & 7 \end{pmatrix} \text{ are } 1,2,3$$

Find eigenvector corresponding to λ=1 and make its norm as Y+1

Your answer: $(\text{----},\text{-------},\text{-------})^T$

20

$$Eigenvalues\ \text{of A=}\begin{pmatrix} 3 & -3 & 4 \\ 2 & -4 & 8 \\ 2 & -4 & 7 \end{pmatrix}\ \text{are 1,2,3}$$

$$What\ \ \text{is Rank of the matrix}\ (\text{A-1}\times\text{I})$$

$$Your\ \text{answer: Rank is ----------}$$

21

$$Eigenvalues \text{ of A=}\begin{pmatrix} 3 & -3 & 4 \\ 2 & -4 & 8 \\ 2 & -4 & 7 \end{pmatrix} \text{ are } 1,2,3$$

$What$ is Rank of the matrix (A-4I)

$Your$ answer: Rank is ----------

22

$$\textit{Eigenvalues} \text{ of } A=\begin{pmatrix} 3 & -3 & 4 \\ 2 & -4 & 8 \\ 2 & -4 & 7 \end{pmatrix} \text{ are } 1,2,3$$

What is Rank of the matrix $(A\text{-}I)(A\text{-}2I)$

Your answer: Rank is ----------

23

$$\textit{Eigenvalues} \text{ of } A=\begin{pmatrix} 3 & -3 & 4 \\ 2 & -4 & 8 \\ 2 & -4 & 7 \end{pmatrix} \text{ are } 1,2,3$$

What is Rank of the matrix $(A\text{-}I)(A\text{-}2I)(A\text{-}3I)$

Your answer: Rank is ----------

24

Let $\lambda_1, \lambda_2, \lambda_3$ are three distint eigenvalues of 3×3 matrix

(*Later* we will deal the case of repeated eigenvalues)

Let x_1, x_2, x_3 be corresponding eigenvectors.

$$A\begin{matrix}|\\x_1\\|\end{matrix} = \lambda_1 \begin{matrix}|\\x_1\\|\end{matrix}$$

$$A\begin{matrix}|\\x_2\\|\end{matrix} = \lambda_2 \begin{matrix}|\\x_2\\|\end{matrix}$$

$$A\begin{matrix}|\\x_3\\|\end{matrix} = \lambda_3 \begin{matrix}|\\x_3\\|\end{matrix}$$

$$A\underbrace{\begin{pmatrix}| & | & |\\ x_1 & x_2 & x_3\\ | & | & |\end{pmatrix}}_{S} = \begin{pmatrix}| & | & |\\ \lambda_1 x_1 & \lambda_2 x_2 & \lambda_3 x_3\\ | & | & |\end{pmatrix} = \underbrace{\begin{pmatrix}| & | & |\\ x_1 & x_2 & x_3\\ | & | & |\end{pmatrix}}_{S}\underbrace{\begin{pmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3\end{pmatrix}}_{\Lambda}$$

$AS = S\Lambda$; Later we will prove that S is invertble

$A = S\Lambda S^{-1}$

What is A^{100}

Your Answer in terms of S and Λ is $\mathrm{A}^{100} = ---$

25

$$A = \begin{pmatrix} Y & Y+1 \\ Y+2 & Y+4 \end{pmatrix}$$

$Trace\left(A^T A\right)$ in terms of Y is ---

26

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & Y \end{pmatrix}$$

Trace of AA^T is -----

27

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & Y \end{pmatrix}$$

Square of $\mathit{Frobenius}$ norm of A is -----

28

Square of *Frobenius* norm of A is same as Trace of matrices ----- and ------

29

Rank of matrix A is same as that of matrices ----- and ------

30

$Rank$ of 5×5 Matrix A is 2.

Then at least --- (number of) eigenvalues are -----(value)

Do not go for Revolution
(Avoid studying only before examination)
Go for Evolution