Week -2

- **Rank** of a matrix refers to no. of linearly independent rows or columns of the matrix.

- The **null space** of a matrix A consists of all vectors β such that $A\beta = 0$ and $\beta \neq 0$.

- **Nullity** of a matrix is the no. of vectors in the null space of the given matrix

- The size of the null space ~~&~~ provides us with the no. of linear relations among the attributes

Nullity of A + Rank of A = Total no. of attributes of A

(No. of equations) (No. of independent variables) (Total no. of variables)

→ Moore penrose pseudo inverse of a matrix -

$$Ax = b$$
$$x = A^{+}b$$

A^{+} → pseudo inverse or generalised inverse

In R, "ginv(A)" is used to calculate this generalised inverse.

$Ax = b$

- $m = n$
  - A is full rank and $|A| \neq 0$ — $x = A^{-1}b$ — Unique solution
  - A is not full rank
    - consistent (∞ solutions)
    - Inconsistent (no solution)
- $m > n$ — $x = A^{+}b$ (Least square solution)
- $m < n$ — $x = A^{+}b$ (Least norm solution)

where we define pseudo inverse A^{+} appropriately

Vectors & Distances $\rightarrow$

$$x^1 = \begin{bmatrix} x_1^1 \\ x_2^1 \end{bmatrix} \quad x^2 = \begin{bmatrix} x_1^2 \\ x_2^2 \end{bmatrix}$$

Distance b/w these two can be calculated as -

$$l = \sqrt{(x_1^2 - x_1^1)^2 + (x_2^2 - x_2^1)^2}$$

$$l = \sqrt{(X_2 - X_1)^T (X_2 - X_1)}$$

Scanned with CamScanner

Unit vector → is a vector with magnitude 1 (distance from origin)

Unit vectors are used to define directions in a coordinate system.

$$\hat{a} = \frac{A}{|A|}$$

Orthogonal vectors →
Two vectors are orthogonal to each other when their dot product is zero.

Dot product (scaler product) of two n dimensional vectors A & B →

$$A \cdot B = \sum_{i=1}^{n} a_i b_i$$

Thus the vectors A & B are orthogonal to each other if and only if

$$A \cdot B = \sum_{i=1}^{n} a_i b_i = A^T B = 0$$

Orthonormal vectors → are orthogonal vectors with unit magnitude

All orthonormal vectors are orthogonal.

Basis vectors → are set of vectors that are independent and span the space

Eg: two vectors $v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Can span R^2 and are independent and hence form the basis for the R^2 space.

Basis vectors are NOT unique

If $a^T b = 0$, then a and b are orthogonal

Representation of line and plane →

$$n = \begin{bmatrix} a \\ b \end{bmatrix}$$

$$ax_1 + bx_2 + c = 0$$

$$n^T X + b = 0$$

In three dimensions,

- One equation would represent a plane
- Two equations would represent a line

because we have 3 degrees of freedom; if we take away 2, then we are left with 1. and a one-dimensional object is a line

- Three equations would represent a point (3 - 3 = 0 degree of freedom)

Projections →

We can define the projection $(\hat{x})$ of a vector (x) onto a lower dimension (two dimensions in the picture) mathematically as

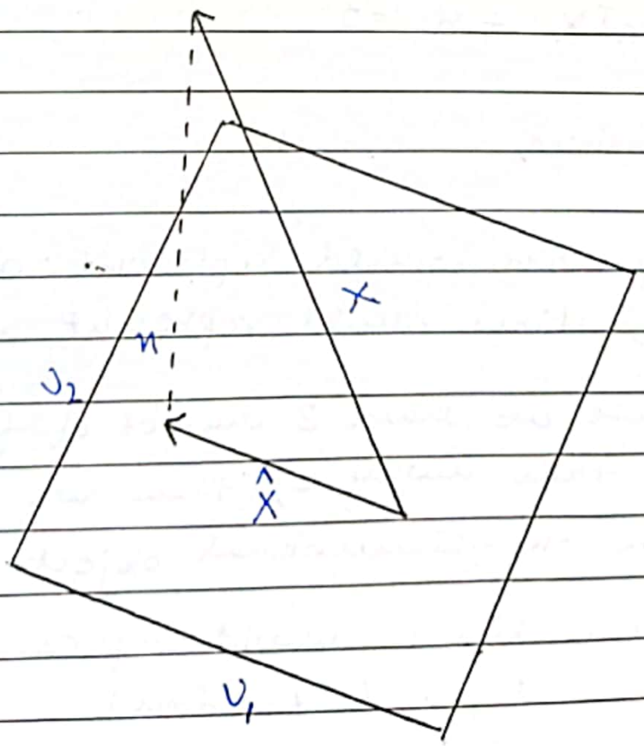
$$\hat{x} = c_1 v_1 + c_2 v_2$$

Using vector addition,

$$x = c_1 v_1 + c_2 v_2 + n$$

→ Since a plane is a two-dimensional object, we will need two basis vectors, let them be v_1 & v_2

So, any line on the plane can be written as a combination of v_1 & v_2



Since n is perpendicular to the plane, $n^T v_1 = v_1^T n = 0$
And similarly, $n^T v_2 = v_2^T n = 0$

Eg : $x = [1\,2\,3]^T$

Projecting this vector onto space spanned by the vectors $v_1 = [1\ -1\ -2]^T$ and $v_2 = [2\ 0\ 1]^T$

Thus, finding the projection onto the plane defined by v_1 & v_2 is

$$\hat{x} = \underbrace{\frac{x^T v_1}{v_1^T v_1}}_{c_1} v_1 + \underbrace{\frac{x^T v_2}{v_2^T v_2}}_{c_2} v_2$$

$$\hat{x} = \frac{[1\,2\,3]\begin{bmatrix}1\\-1\\-2\end{bmatrix}}{6}\begin{bmatrix}1\\-1\\-2\end{bmatrix} + \frac{[1\,2\,3]\begin{bmatrix}2\\0\\1\end{bmatrix}}{5}\begin{bmatrix}2\\0\\1\end{bmatrix}$$

$$\hat{x} = -\frac{7}{6}\begin{bmatrix}1\\-1\\-2\end{bmatrix} + 1\begin{bmatrix}2\\0\\1\end{bmatrix}$$

$$\hat{x} = \begin{bmatrix}5/6\\7/6\\20/6\end{bmatrix}$$

Projection - Generalisation →

Projections onto general directions

⇒ Projection of X onto a space spanned K linearly independent vector

$$\hat{x} = \sum_{j=1}^{K} c_j v_j$$

$$\hat{x} = \begin{bmatrix} v_1 & \cdots & v_K \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_K \end{bmatrix}$$

$$\hat{x} = Vc$$

Using orthogonality idea

$$V^T(x - \hat{x}) = V^T(x - Vc) = 0$$

$$V^T x - V^T V c = 0$$

$$c = (V^T V)^{-1} V^T x$$

$$\hat{x} = V(V^T V)^{-1} V^T x$$

Hyperplanes →

Hyperplane is a geometric entity whose dimension is one less than that of its ambient space.

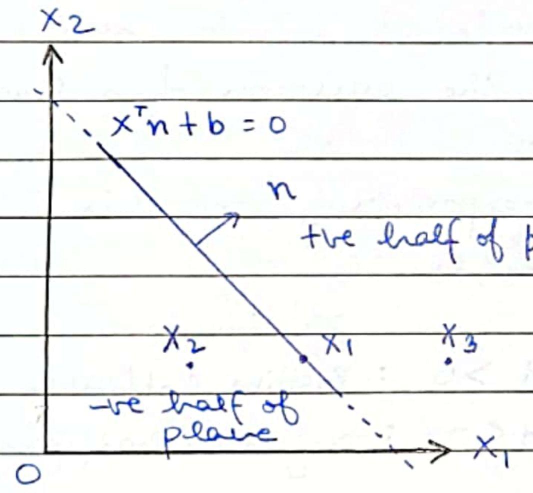
Eg : the hyperplanes for a 3D space ~~is~~

are 2D planes and hyperplanes for a 2D space are 1D lines & so on.

The equation for a hyperplane is $x^T n + b = 0$
for 2D, it would be $x_1 n_1 + x_2 n_2 + b = 0$

If the plane goes through the origin, then the hyperplane also becomes a subspace

Half space →



In above graph, whole two-dimensional space is broken into two spaces. These two spaces are called half-spaces.

$$x^T n + b = 0 \text{ for every } x \in \text{line}$$

$x^T n + b > 0$ for every $x \in$ subspace in the n-direction (x_3)

$x^T n + b < 0$ for every $x \in$ subspace in the '-n' direction (x_2)

Scanned with CamScanner

Let us consider a 2D geometry with $n = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$
and $b = 4$

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$x^T n + b = 0$$

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} + 4 = 0$$

$$x_1 + 3x_2 + 4 = 0$$

The hyperplane is the equation of a line

The halfspaces corresponding to this hyperplane are

$x_1 + 3x_2 + 4 > 0$: Positive halfspace
$x_1 + 3x_2 + 4 < 0$: Negative halfspace

Eigenvalues and Eigenvectors →

Eigenvalues are the special set of scaler values that is associated with the set of linear equations most probably in the matrix equations.

Eigenvalue is a scaler that is used to transform the eigenvector.

The basic equation is $Ax = \lambda x$

The number or scaler value 'λ' is an eigenvalue of A.

An eigenvector corresponds to the real non-zero eigenvalues which point in the direction stretched by the transformation
WHEREAS
Eigenvalue is considered as a factor by which it is stretched.
In case if the eigenvalue is negative, the direction of the transformation is negative

Eigenvectors are the vectors (non-zero) that do NOT change the direction when any linear transformation is applied. It changes only by a scaler factor

In $Ax = \lambda x$

x is an eigenvector of A corresponding to eigenvalue, λ.

Eigenvalue of Square Matrix →

Let $A_{n \times n}$ is a square matrix

Then $[A - \lambda I]$ is called an eigen matrix, which is an indefinite or undefined scaler.

The determinant of eigen matrix can be written as $|A - \lambda I|$.

Here, $|A-\lambda I| = 0$ is the Eigen equation where I is the identity matrix.

Connections b/w eigenvectors, column space and null space →

Eigenvalues can be complex even for real matrices.

When eigenvalues become complex, eigenvectors also become complex

However, if the matrix is symmetric, then the eigenvalues are always real

As a result, eigenvectors of symmetric matrices are also real.

Further, there will always be 'n' linearly independent eigenvectors for symmetric matrices.

→ Eigenvalues of matrices of the form A^TA or AA^T while ~~boig~~ being real, are also non-negative.

$$A \overset{\text{Eigenvector}}{v} = \underset{\text{Eigenvalue}}{\lambda} v$$

When eigenvalue becomes zero $(Av = 0)$

Then the eigenvectors corresponding to zero eigenvalues are in the null space of the matrix.

Conversely, if the eigenvalue corresponding to an eigenvector is not zero, then the eigenvector can not be in the null space

- Lets assume $A_{n \times n}$ to be a symmetric matrix.
- There are 'r' zero eigenvalues.
- So, there will be 'n-r' non-zero eigenvalues.
- The 'r' eigenvectors corresponding to 'r' zero eigenvalues are all in the null-space
- So, the dimension of the null-space is 'r'.

From Rank-nullity theorem,

$$\text{Rank} + \underset{(r)}{\text{Nullity}} = \text{No. of columns } (n)$$

$$\text{Rank} = n - r$$

And as column rank = row rank

So, both column rank & row rank are equal to 'n-r'.

⇒ There are 'n-r' independent vectors in the columns of the matrix.

→ Dimension of column space is 'n-r' (from rank-nullity theorem)

⇒ The eigenvectors corresponding to the ~~zero~~ non-zero eigenvalues form a basis for the column space.

→ The eigenvectors corresponding to zero eigenvalues can be used to identify relationships among variables.