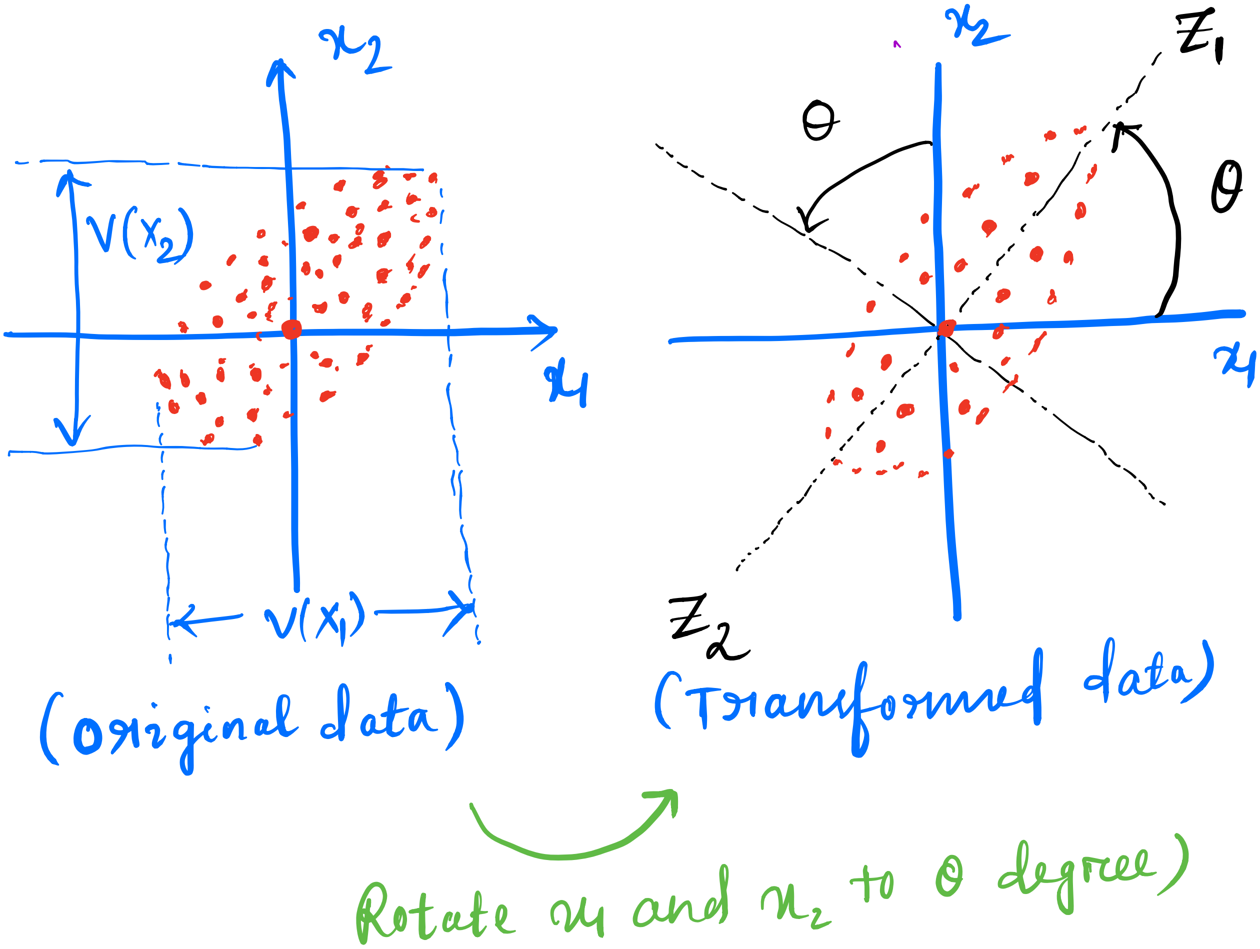
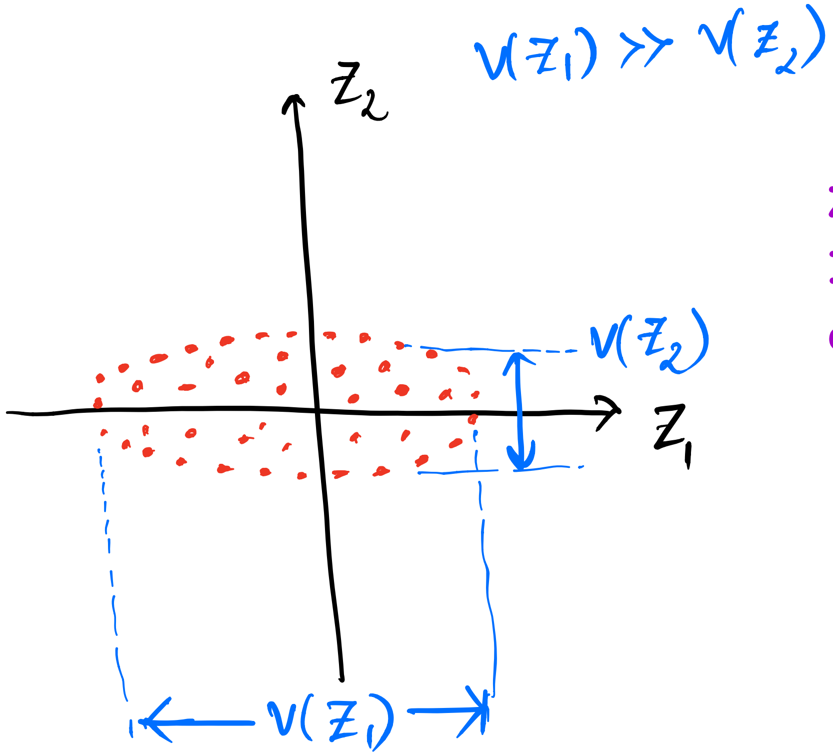
PCA

Principal Component Analysis



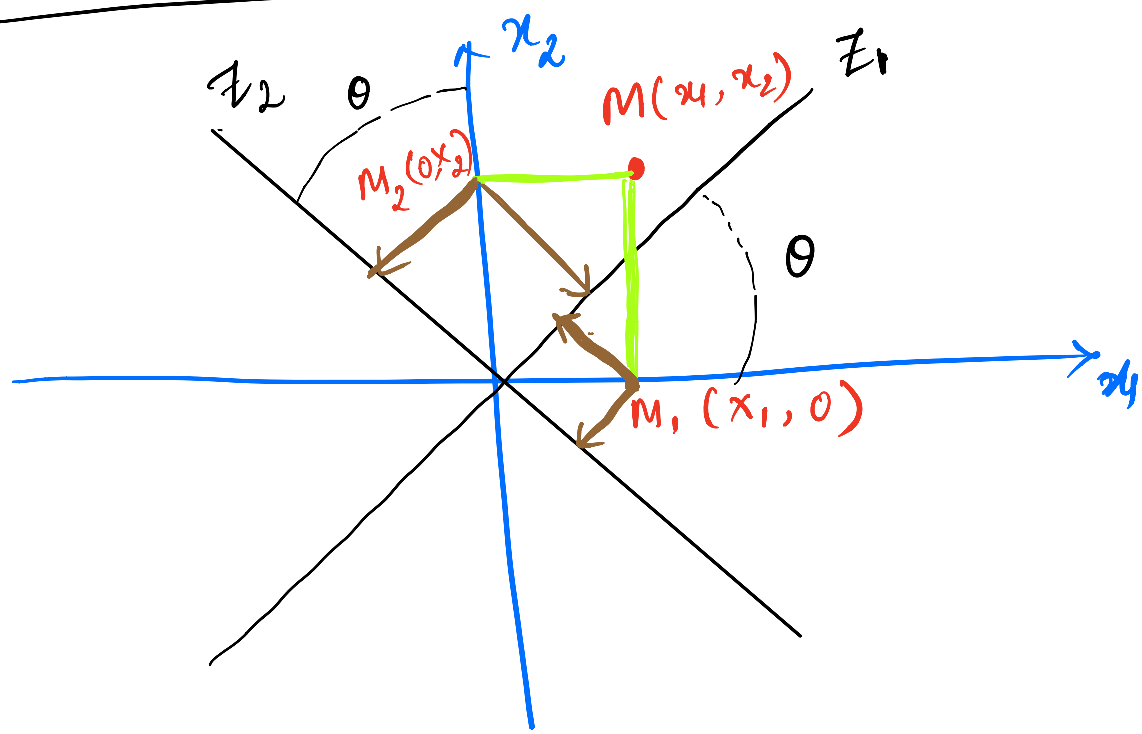


Z_1 and Z_2 are Independent of each other.

↓

Orthogonal to each other.

How to Project a point from x_1, x_2 to Z_1 Z_2



In Real time some columns are important and some columns are not.

X	Y	Z
1	1	1
0.5	0	0
0.25	1	1
0.35	1.5	1.5
0.45	1	1
0.57	2	2.1
0.62	1.1	1

Is Z- is adding any new Information beyond what is already contained in y?

Y and Z are highly correlated.

Derivation

$$\begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix}_{2\times 2} \times \begin{bmatrix} 3 \\ 2 \end{bmatrix}_{2\times 1} = \begin{bmatrix} 12 \\ 8 \end{bmatrix}_{2\times 1} = 4 \times \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Input matrix(A) Eigen vector(v) eigen value(λ) eigen vector(v)

$$\boxed{Av = \lambda \cdot v}$$

Calculating Eigen values: $A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$

$$Av = \lambda v$$
$$\Rightarrow Av - \lambda v = 0$$
$$\Rightarrow (A - \lambda) = 0$$
$\Rightarrow (A - \lambda I) = 0$, where I is the identity matrix.

Then, $|A - \lambda I| = \left| \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right|$

$$= \left| \begin{bmatrix} -\lambda & 1 \\ -2 & -3-\lambda \end{bmatrix} \right|$$

$$= (-\lambda \times (-3 - \lambda))$$

$$= \lambda^2 + 3\lambda + 2$$

$$\boxed{\lambda_1 = -1 \quad , \quad \lambda_2 = -2}$$ $\rightarrow$ Eigen values

$\boxed{\text{Step-1}}$

↓

find co-variance matrix

$$\text{Variance } (s^2) = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}$$

$$\text{Covariance, } \operatorname{cov}(x,y) = \sum_{i=1}^{n} \frac{(x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

e.g: covariance matrix of 3 rows & 3 columns:

covariance of itself is nothing but variance

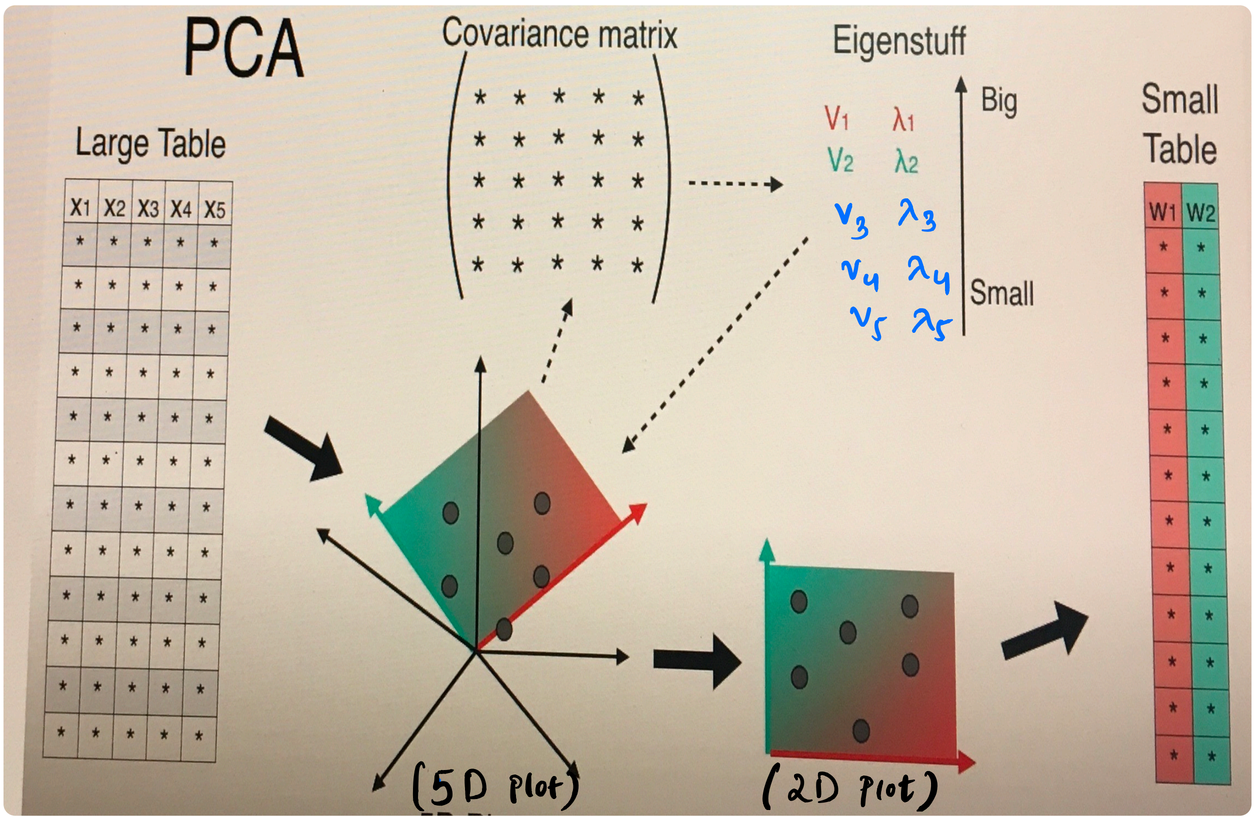
$$\begin{bmatrix} \operatorname{cov}(x,x) & \operatorname{cov}(x,y) & \operatorname{cov}(x,z) \\ \operatorname{cov}(y,x) & \operatorname{cov}(y,y) & \operatorname{cov}(y,z) \\ \operatorname{cov}(z,x) & \operatorname{cov}(z,y) & \operatorname{cov}(z,z) \end{bmatrix}$$

$\boxed{\text{Step-2}}$

Calculate the eigen vectors and eigen values of the covariance matrix.

$\boxed{\text{Step-3}}$

Select m eigen vectors that corresponds to the m largest eigen values to be the new axis.



PCA is an orthogonal linear transformation that transfers the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie in the 1st coordinate, the second greatest variance lies in the 2nd coordinates and so on - - - -