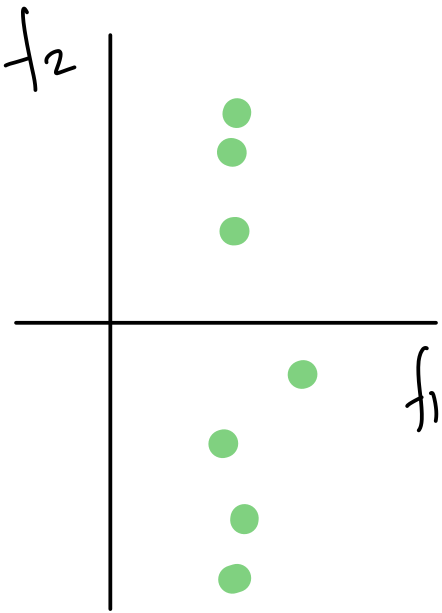
PCA

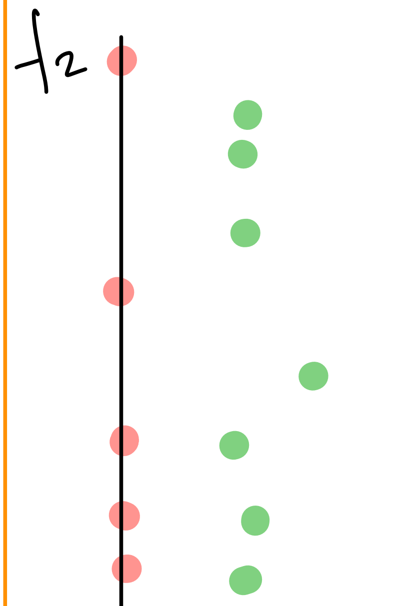
PCA

Principal Component Analysis.

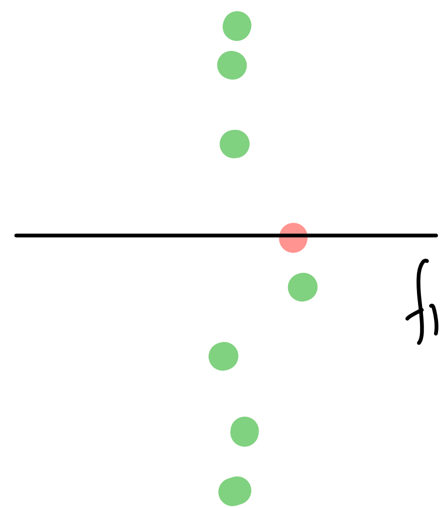
Dimensionality Reduction Technique.

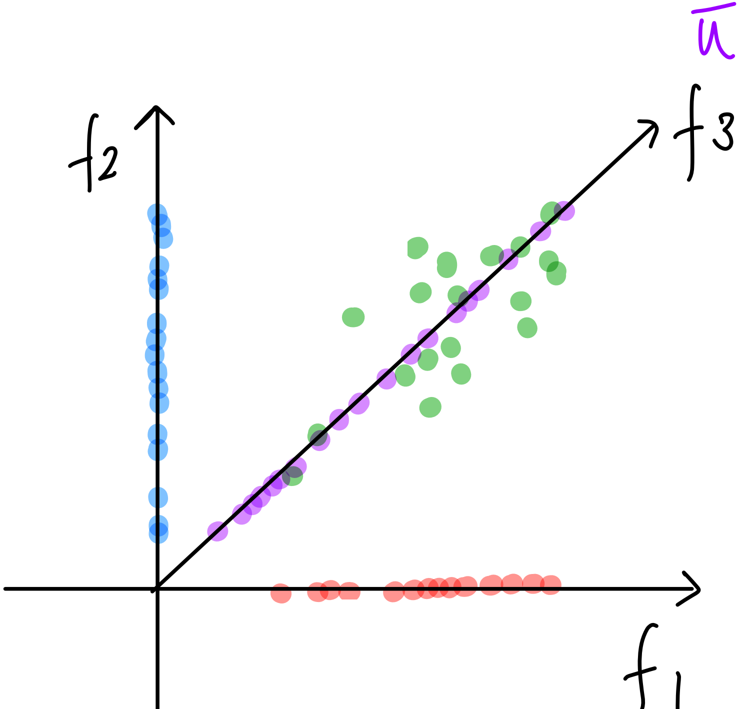


(I)

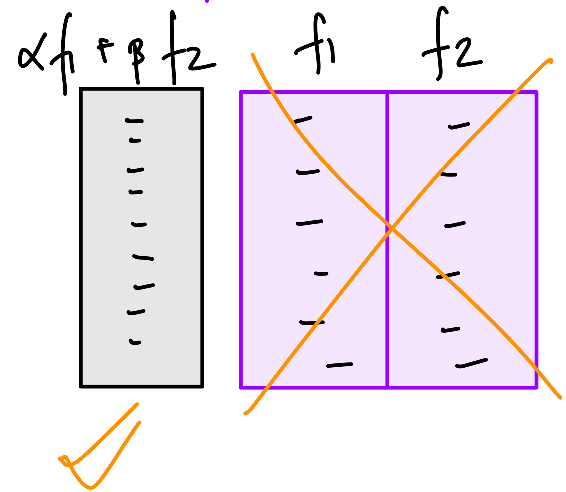


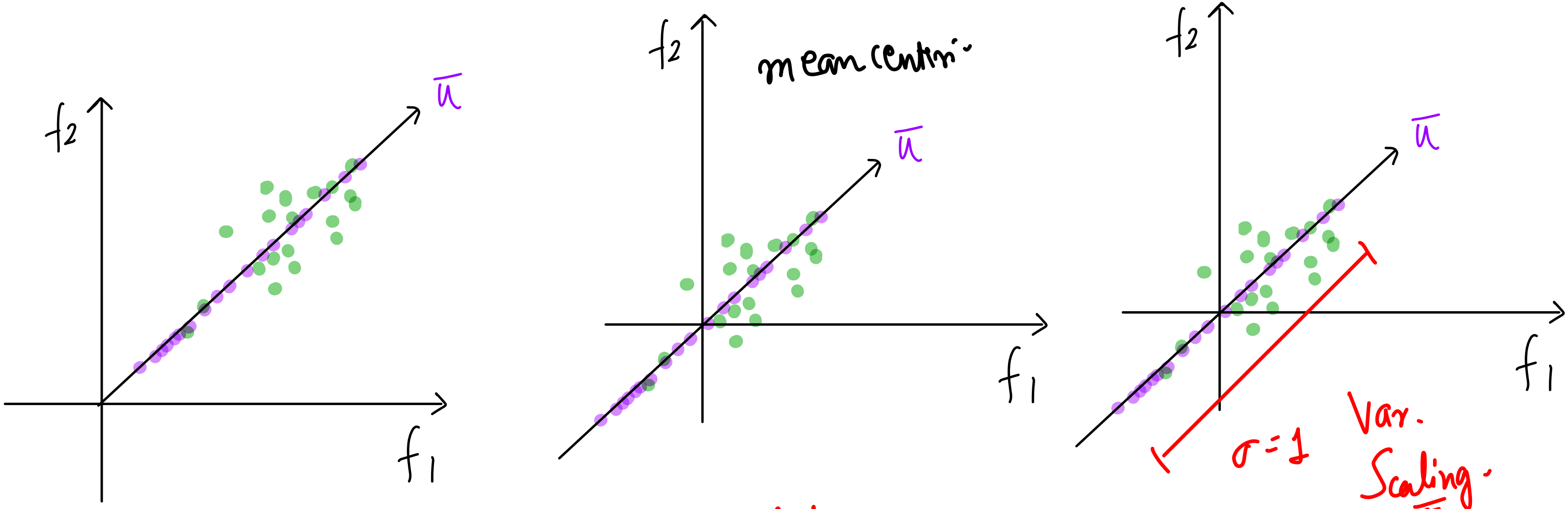
(II)





$U = \alpha f_1 + \beta f_2$





Standardise the data.

$$\left(\frac{x - \mu}{\sigma}\right) =$$

mean

std

0

$\mu = 0 \leftarrow$ mean centering

$\sigma = 1 \leftarrow$ Variance Scaling

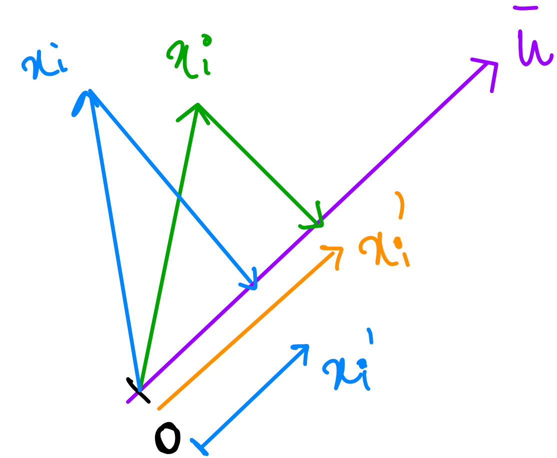
$\underbrace{68}_{1\sigma} - \underbrace{95}_{2\sigma} - 99$

$-3\sigma \quad 3\sigma$

$-\sigma \quad \sigma$

$-2\sigma \quad 2\sigma$

$$x_i' = x_i \cdot \hat{U} = \frac{x_i \cdot \bar{U}}{\|U\|}$$



<u>OBJECTIVE</u>. To find out $\bar{u}$ when x_i's are projected, the variance should be maximum.

$$x_i' = x_i \hat{U} = \frac{x_i U}{\|u\|}$$

$$\Rightarrow x_i U \quad \text{st} \quad \|U\| = 1$$

Maximise the variance of x_i' on $\bar{u}$.

$$\max_{u} \{ \text{Var. of } x_i' \} \text{ such that } \|u\| = 1$$

$$x_i' = \underbrace{u^T}_{\text{vector}} \, x_i \quad \text{(}x_i \to \text{point)}$$

$$u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ \vdots \end{bmatrix} \qquad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_n \\ \vdots \end{bmatrix}$$

$$\mathrm{Var}(x_i') = \frac{1}{n} \sum \left(x_i' - \cancel{\bar{x}_i'} \right)^2 \to 0$$

$$\bar{x}_i' = \frac{x_1' + x_2' + x_3' + x_4' \dots}{n}$$

Mean of projection of data on $\bar{u}$

$$\bar{x}_i' = \frac{1}{n}\sum_{i=1}^{n} x_i' = \frac{1}{n}\sum u^T x_i$$

$$\max_{u} \ \frac{1}{n}\sum \left(x_i' - \cancelto{0}{\bar{x}_i'}\right)^2 \quad \text{s.t} \quad \|u\| = 1$$

$$\max_{u} \ \frac{1}{n}\sum \left(x_i'\right)^2 \quad \text{s.t.} \quad \|u\|^2 = 1$$

$$u^T = \begin{bmatrix} u_1 & u_2 & u_3 & u_4 \end{bmatrix} \quad u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \Rightarrow u^T u = u_1^2 + u_2^2 + u_3^2 + u_4^2$$

$$\max_{u} \frac{1}{n} \sum (x_i')^2 \quad s.t \quad u^T u = 1$$

$$\boxed{\max_{u} \frac{1}{n} \sum (u^T x_i)^2 \quad s.t \quad u^T u = 1}$$

GD

Matrix theory ⇒ Matrix Multiplication

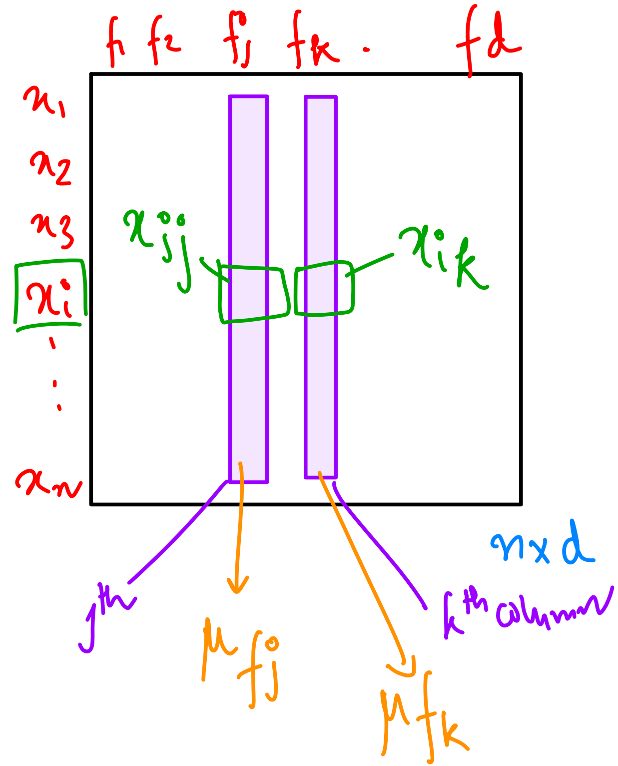
$$\max_{u} \frac{1}{n} \sum (u^T x_i)^2 + \lambda (u^T u - 1)$$

Original data → Standardised data → projection of x_i on $\bar{u}$

x_i $\quad$ x_i $\quad$ x_i' $\quad$ x_i

$$S_{jk} = Cov(f_j, f_k)$$

$$= \frac{1}{n} \sum \Big(x_{ij} - \underbrace{\cancelto{0}{\bar{x}_{ij}}}_{\mu_{f_j}}\Big)\Big(x_{ik} - \underbrace{\cancelto{0}{\bar{x}_{ik}}}_{\mu_{f_k}}\Big)$$



$$\boxed{S_{jk} = \frac{1}{n} \sum \big(x_{ij} \cdot x_{ik}\big)}$$

$$Cov(x, y) = \frac{1}{n} \sum (x - \bar{x})(y - \bar{y})$$

$$S = \begin{array}{c|cccc} & f_1 & f_2 & f_k & \cdots\ f_d \\ \hline f_1 & \mathrm{Cov}(f_1,f_1) & & & \mathrm{Cov}(f_1,f_d) \\ f_2 & \mathrm{Cov}(f_2,f_1) & & & \\ f_k & \mathrm{Cov}(f_k,f_1) & & \mathrm{Cov}(f_k,f_k) & \\ \vdots & & & & \\ f_d & \mathrm{Cov}(f_d,f_1) & & & \mathrm{Cov}(f_d,f_d) \end{array}$$

$d \times d$

S^T

SQUARE SYMMETRIC MATRIX

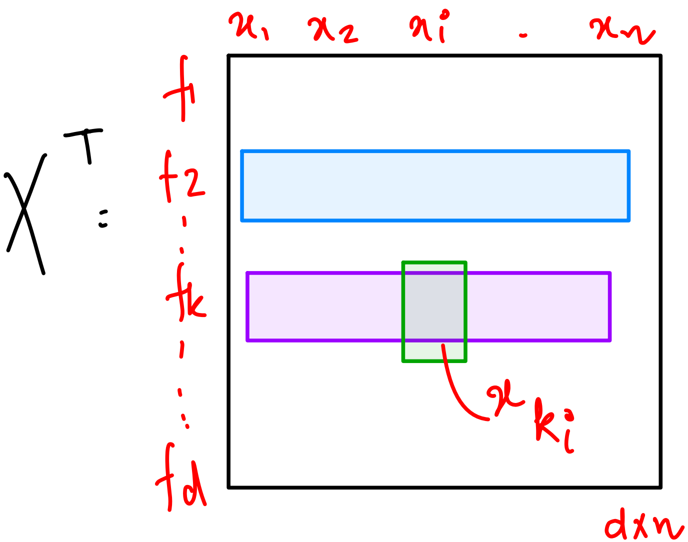
$$\boxed{S = S^T}$$

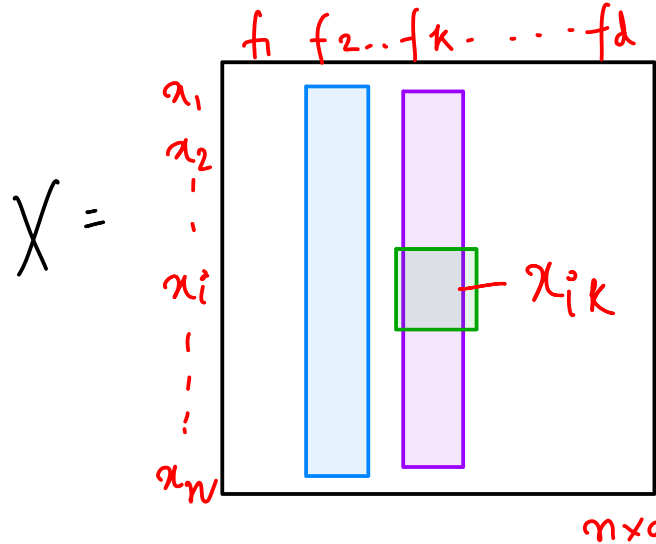
$\mathrm{Cov}(x, y)$ ①

$$= \frac{1}{n} \sum (x - \bar{x})(y - \bar{y})$$

$\mathrm{Cov}(y, x)$ ②

$$= \frac{1}{n} \sum (y - \bar{y})(x - \bar{x})$$





$$S_{d\times d} = \frac{1}{n} \underset{d\times n}{X^T} \underset{n\times d}{X}$$

$$S_{kj} = \frac{1}{n} \sum x_{ik} \cdot x_{ij}$$

$$\frac{1}{n}\begin{bmatrix} f_1f_1 & f_1f_2 & f_1f_k & f_1f_d \\ f_2f_1 & & & \\ f_3f_1 & & f_kf_n & \\ f_df_1 & & & f_df_d \end{bmatrix}$$

$$\sum_{i=1}^{n} \left(U^T x_i\right)^2$$

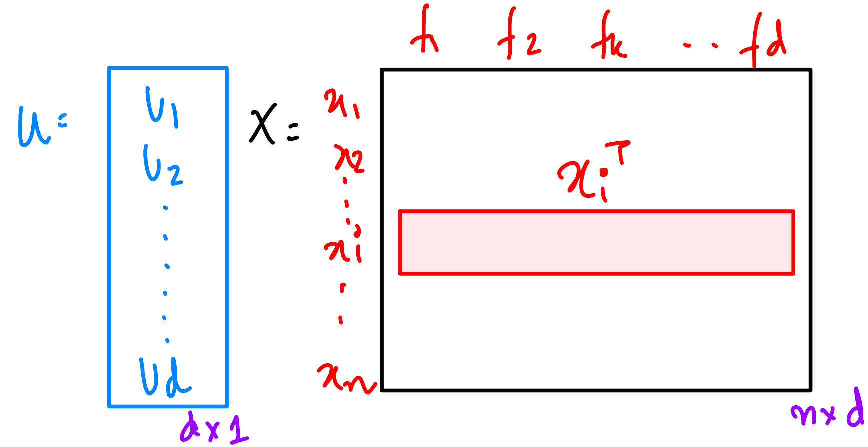
Scaler

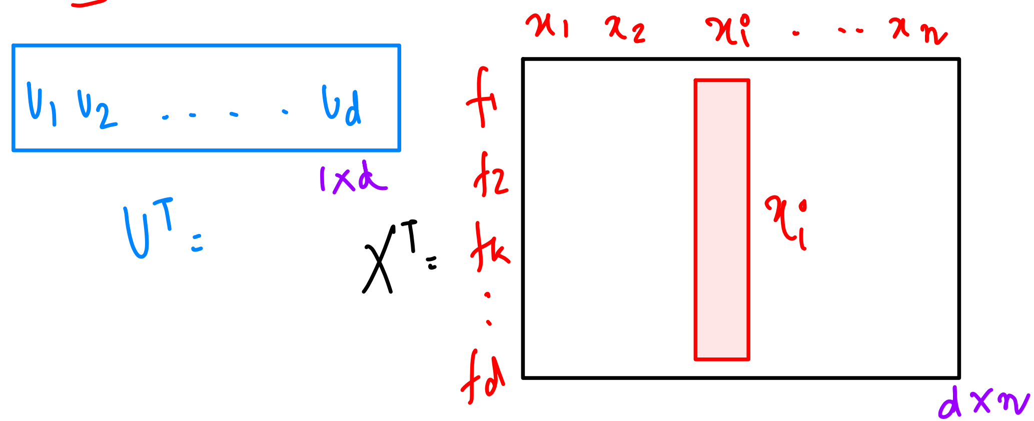
$$\underset{1\times d}{U^T}\ \underset{d\times n}{X^T} = \underset{1\times n}{A}$$

$$= \left[U^T x_1 \quad U^T x_2 \quad U^T x_3 \quad U^T x_i \quad U^T x_n\right]$$

$$A = U^T x_i$$

$$\boxed{\underset{1\times n}{A} = \underset{1\times d}{U^T}\ \underset{d\times n}{X^T}}$$





$$\underset{n\times 1}{B} = \underset{n\times d}{X}\;\underset{d\times 1}{U}$$

$$\underset{n\times 1}{B} = x_i^T U$$

$$X = \begin{array}{c} \\ x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{array} \begin{array}{c} \begin{array}{ccccc} f_1 & f_2 & f_k & \cdots & f_d \end{array} \\ \left[\begin{array}{c} \\ \\ x_i^T \\ \boxed{\qquad\qquad\qquad} \\ \\ \\ \end{array}\right] \end{array}_{n\times d} \qquad U = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_d \end{bmatrix}_{d\times 1}$$

$$B = \begin{bmatrix} x_1^T u \\ x_2^T u \\ x_i^T u \\ \\ x_n^T u \end{bmatrix}_{n\times 1}$$

$$\underline{\underline{Res}} \Rightarrow \frac{1}{n}\sum (u^T x_i)^2 = \frac{1}{n}\sum \underbrace{(u^T x_i)}_{A}\underbrace{(u^T x_i)^T}_{x_i^T u^{T^T}}$$

$$\boxed{x_i^T u} \; B$$

$$\underline{\underline{Res}} = \frac{1}{n} A_{1\times n} \cdot B_{n\times 1}$$

$$= \frac{1}{n} U^T X^T \cdot X U$$

$$\max_u \frac{1}{n}\sum (u^T x_i)^2 \quad \text{s.t.} \quad (\|u\|^2 = 1)$$

$$\max_u \frac{1}{n} (U^T X^T X U) \quad \text{s.t.} \quad U^T U = 1$$

$$\min_{u} \; -\frac{1}{n}\left(u^T \underbrace{X^T X}_{} u\right) + \lambda\left(u^T u - 1\right)$$

$$S = \frac{1}{n} X^T X$$

↳ lagranges multiplier.

$$\min_{u} \; \underbrace{-\left(u^T S u\right)}_{A} + \lambda\left(u^T u - 1\right)$$

$$\frac{\partial}{\partial u}\left[-\overbrace{u^T \underbrace{S u}_{}}^{}\right] = -\left(u^T S\right)^T = -S^T u - S u$$

$$= -2 S u$$

$$\frac{\partial}{\partial u}\left[\lambda\, u^T u\right] = 2 \lambda u$$

$$\frac{\partial A u}{\partial u} = A^T$$

$$\frac{\partial\, u^T A}{\partial u} = A$$

$$\frac{\partial}{\partial u} u^T u = 2u$$

$$-2SU + 2\lambda U = 0$$

$$\underbrace{S}_{\text{matrix}}\ \underbrace{U}_{\text{vector}} = \underbrace{\lambda}_{\text{constant}}\ U \rightarrow \text{vector}.$$

$$\underset{\text{matrix}}{S}\ \underset{\text{Eigen Vector}}{U} = \underset{\text{Eigen value}}{\lambda}\ \underset{\text{Eigen vector}}{U}$$

1. Dataset ⇒ Standardised.

2. Compute $S_{d \times d}$ ⇒ Covariance Matrix of X

3. Eigen Vector & Eigen Values.

$u_1 \quad u_2 \quad u_3 \quad u_4 \quad u_5$

$\lambda_1 \quad \lambda_2 \quad \lambda_3 \quad \lambda_4 \quad \lambda_5$ } variances

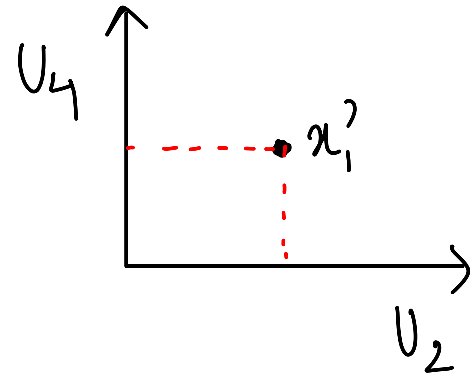
0.2 0.3 0.1 0.4 0.25

$\lambda_4 > \lambda_2 > \lambda_5 > \lambda_1 > \lambda_3$

u_4 ✓ u_2 ✓ u_5 ✗ u_1 ✗ u_3 ✗

x_i^{o}

$$\left.\begin{aligned} x_1 \cdot U_4 &= x_1' \\ x_1 \cdot U_2 &= x_1' \end{aligned}\right\} x_1^{new}$$



$$S = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix}$$

$$\Rightarrow \underbrace{S}_{d \times d} \underbrace{v}_{d \times 1} = \lambda \, v_{d \times 1}$$

$$\Rightarrow S v - \lambda v = 0$$

$$S v - \lambda I v = 0$$

$$(S - \lambda I) v = 0$$

$$\det(S - \lambda I) = 0$$

$S = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix}$ $\qquad$ $S - \lambda I$

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

$$\Rightarrow \det \begin{bmatrix} 1-\lambda & 1 \\ 4 & 1-\lambda \end{bmatrix} =$$

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

$$1^2 + \lambda^2 - 2\lambda - 4 = 0$$

$$\Rightarrow \lambda^2 - 2\lambda - 3 = 0$$

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

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

$$\begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

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

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

$$\boxed{\lambda = -1 \;,\; +3}$$

$\lambda = -1$

$$\begin{bmatrix} 1-\lambda & 1 \\ 4 & 1-\lambda \end{bmatrix} = \begin{bmatrix} 1-(-1) & 1 \\ 4 & 1-(-1) \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} \xrightarrow{R_2 \to R_2 - 2R_1}$$

$$\Downarrow$$

$$\begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}$$

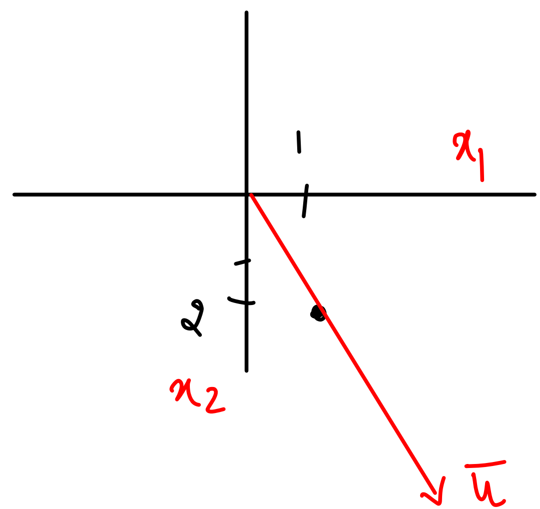
$$\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Rightarrow 2x_1 + x_2 = 0$$

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix} \begin{bmatrix} 2 \\ -4 \end{bmatrix} \begin{bmatrix} 3 \\ -6 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix}, \; 2\begin{bmatrix} 2 \\ -4 \end{bmatrix}, \; 3\begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$\lambda = -1 \Rightarrow$ for $x_1 = 1$, $x_2 = -2$

$$U_{\lambda=-1} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$



$$\lambda = 3$$

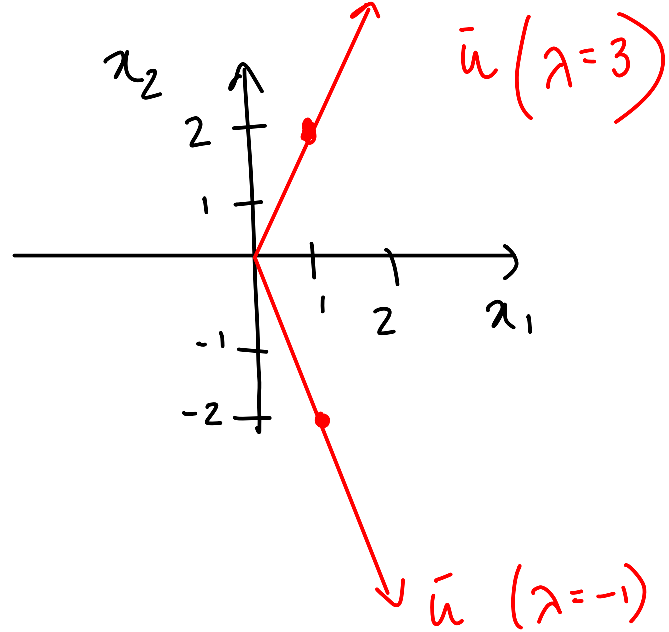
$$\begin{bmatrix} 1-\lambda & 1 \\ 4 & 1-\lambda \end{bmatrix} = \begin{bmatrix} 1-3 & 1 \\ 4 & 1-3 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 4 & -2 \end{bmatrix} \xrightarrow{R_2 \to R_2 + 2R_1}$$

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

1

for $\lambda = 3$, $x_1 = 1$, $x_2 = 2$

$\underline{\underline{\lambda = 3}}$ $\quad v_{\lambda=3} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

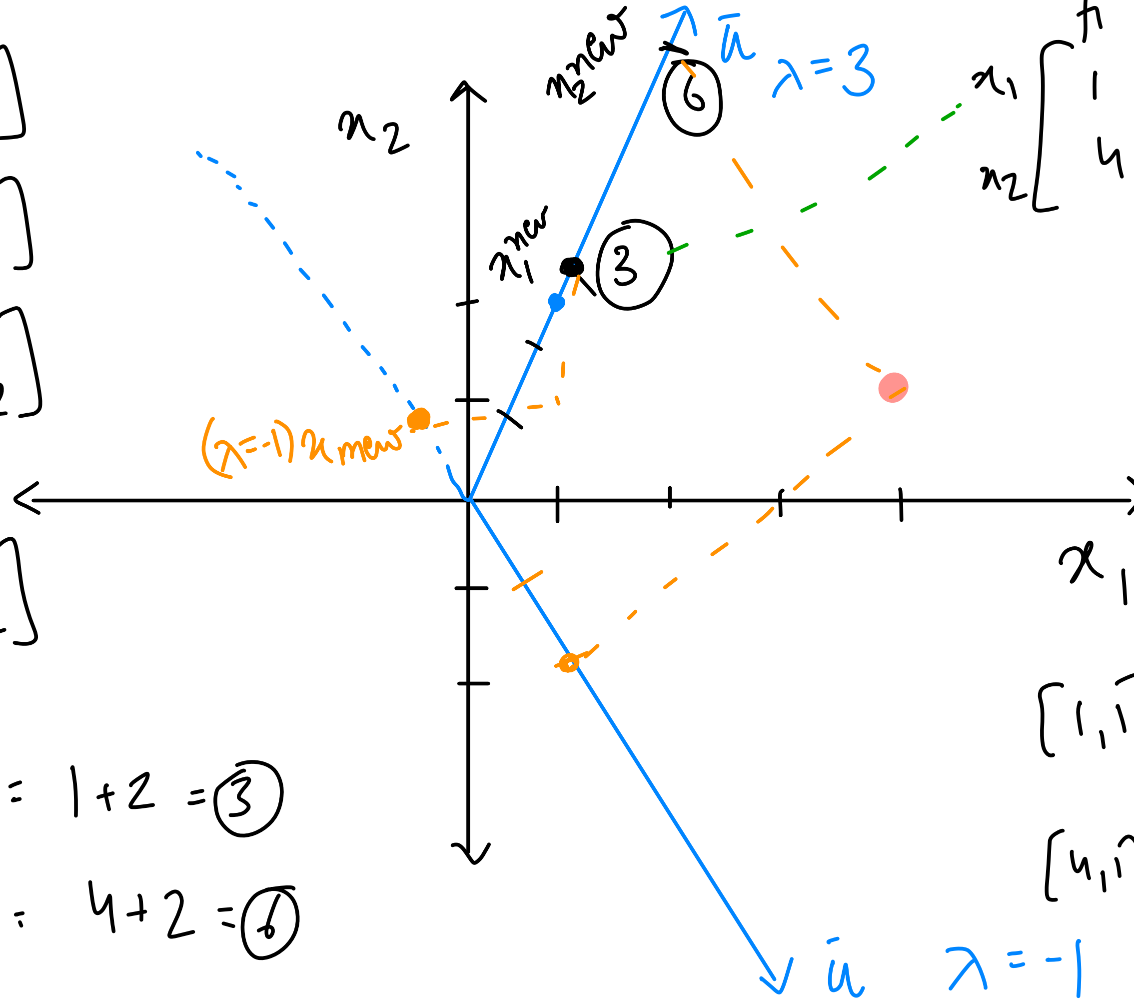


$x_1 = [1, 1]$

$x_2 = [4, 1]$

$v_{\lambda=3} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$v_{\lambda=-1} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$



$$\begin{array}{c|cc} & f_1 & f_2 \\ \hline x_1 & 1 & 1 \\ x_2 & 4 & 1 \end{array}$$

$[1, 1]\begin{bmatrix} 1 \\ 2 \end{bmatrix} = 1 + 2 = 3$

$[4, 1]\begin{bmatrix} 1 \\ 2 \end{bmatrix} = 4 + 2 = 6$

$[1, 1]\begin{bmatrix} 1 \\ -2 \end{bmatrix} = 1 - 2 = -1$

$[4, 1]\begin{bmatrix} 1 \\ -2 \end{bmatrix} = 4 - 2 = 2$