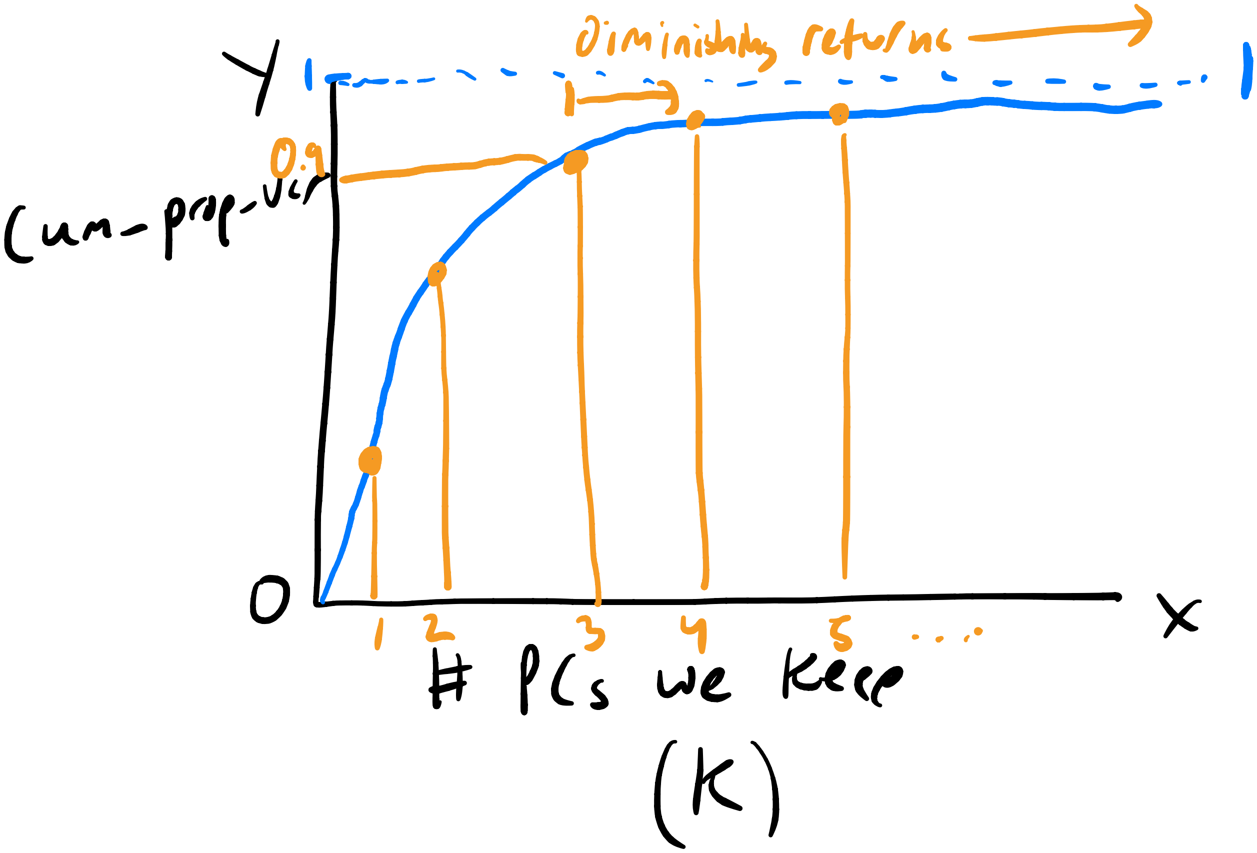
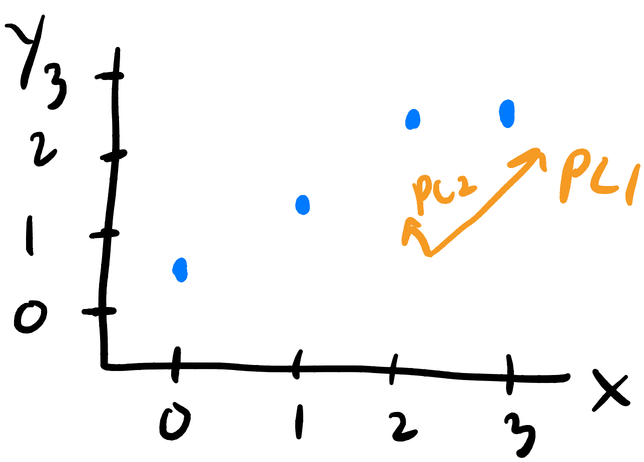


$\Rightarrow$ 90% is a good cut off point: keep PC1, PC2, PC3,
drop PC4 +

Example:

X	Y
0	0.5
1	1.1
2	2.3
3	2.5



1) Normalize → Skip.

2) Center data : $A_c = A - \vec{\mu}$

$\vec{\mu} = (\mu_x, \mu_y)$

$$\mu_x = \frac{0+1+2+3}{4} = 1.5$$

$$\mu_y = \frac{0.5+1.1+2.3+2.5}{4} = 1.6$$

A

x	y
0	0.5
1	1.1
2	2.3
3	2.5

$-[1.5, 1.6] \leftarrow \vec{\mu}$

A_c

x_c	y_c
-1.5	-1.1
-0.5	-0.5
0.5	0.7
1.5	0.9

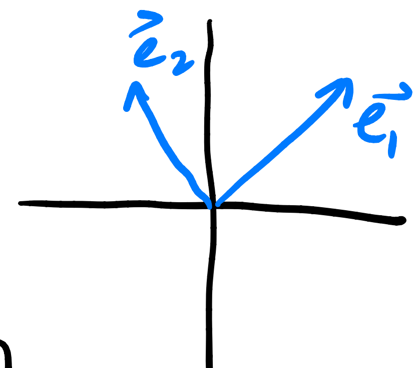
3) Compute $\Sigma = \frac{1}{(N-1)} \cdot (A_c.T) @ (A_c)$

$$\frac{1}{3}\begin{bmatrix} -1.5 & -0.5 & 0.5 & 1.5 \\ -1.1 & -0.5 & 0.7 & 0.9 \end{bmatrix} @ \begin{bmatrix} -1.5 & -1.1 \\ -0.5 & -0.5 \\ 0.5 & 0.7 \\ 1.5 & 0.9 \end{bmatrix}$$

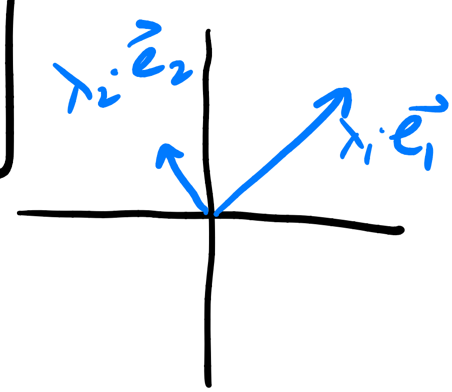
$$= \underbrace{\begin{bmatrix} 1.67 & 1.2 \\ 1.2 & 0.92 \end{bmatrix}}_{\Sigma}$$

4) e_vals, e_vecs of Σ

<u>eigen values</u> : $\left[\overset{\lambda_1}{2.55},\ \overset{\lambda_2}{0.04}\right]$



<u>eigenvectors (P)</u> : $\begin{bmatrix} 0.81 & -0.59 \\ 0.51 & 0.81 \end{bmatrix}$ with columns $\vec{e}_1$, $\vec{e}_2$



5) Sort e_vecs/e_vals high → low according to e_vals

6) a) Compute prop_var by each PC

$$\text{prop\_var} = \left[\underbrace{\frac{2.55}{2.55+0.04}}_{0.98} , \underbrace{\frac{0.04}{2.55+0.04}}_{0.02} \right]$$

PC1: 98% total variance

PC2: 2% total variance

How many PCs we decide to keep

b) Compute cum_prop_var for the top K PCs

$$\text{cum\_prop\_var} = \left[\underbrace{0.98}_{\text{PC1 accounts for 98\% of var}} , \overbrace{\underbrace{0.98 + 0.02}_{\text{PC1 and PC2 account for 100\% variance.}}}^{1.0} \right]$$

c) Define variance threshold to drop PCs.

if drop PC2 → only lose 2% of variance, might be OK.

$$\hat{P} = P[:,0] = \begin{bmatrix} 0.81 \\ 0.59 \end{bmatrix}$$ [if tossing out $\vec{e}_2$].

but here, we will not drop PC2:

$$\hat{P} = \underbrace{P}_{(M,M)}$$ [we won't drop any PCs]

7) project data space → PCA space

$$\overbrace{\hat{A}_c}^{\text{centered data in PCA space}} = \overbrace{A_c}^{\text{centered data}} @ \overbrace{\hat{P}}^{\text{eigenvectors}}$$

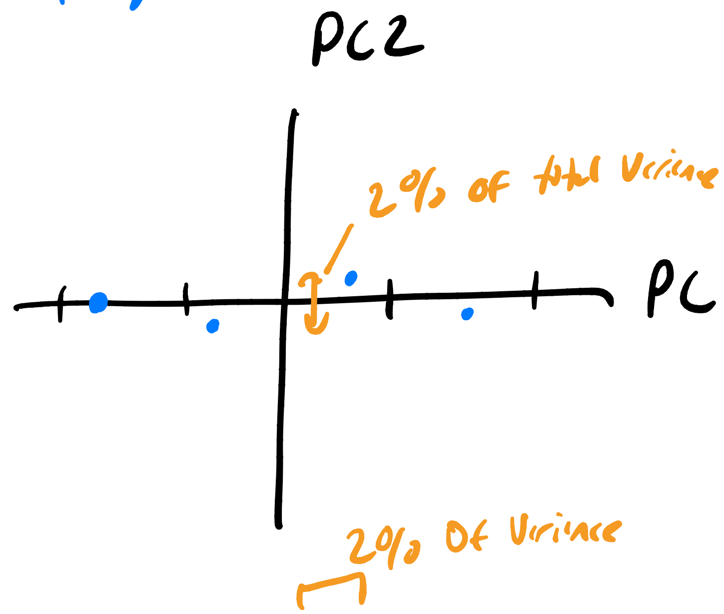
$$\underbrace{\hat{A}_c}_{(N,K)} = \underbrace{A_c}_{(N,M)} @ \underbrace{\hat{P}}_{(M,K)}$$

we are keeping 1st K PCs, dropping rest.

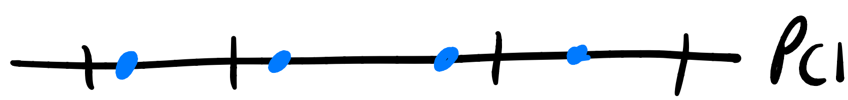
$$\boxed{K \leq M}$$

$$\hat{A}_c = \underbrace{\begin{array}{cc} x & y \\ \left[\begin{array}{cc} -1.5 & -1.1 \\ -0.5 & -0.5 \\ 0.5 & 0.7 \\ 1.5 & 0.9 \end{array}\right] \end{array}}_{(N,M)=(4,2)} \underbrace{\begin{bmatrix} 0.81 & -0.59 \\ 0.59 & 0.81 \end{bmatrix}}_{(M,K)=(2,2)} \qquad K = 2 = M$$

$$\hat{A}_c = \begin{array}{cc} \overbrace{\quad}^{PC1} & \overbrace{\quad}^{PC2} \\ \left[\begin{array}{cc} -1.86 & 0 \\ -0.7 & -0.1 \\ 0.82 & 0.27 \\ 1.74 & -0.16 \end{array}\right] \end{array}$$



if we dropped PC2 :



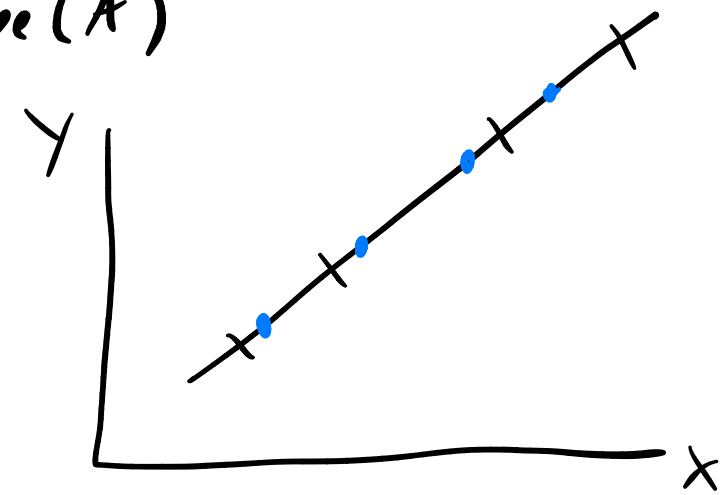
8) Reconstruct data in data space

PCA space → data space

$$\underbrace{A_R}_{(N,M)} = \underbrace{\underbrace{\hat{A}_c}_{(N,K)} @ \underbrace{\overbrace{\hat{P}.T}^{(M,K)}}_{(K,M)}}_{(N,M)} + \underbrace{\vec{\mu}}_{(1,M)}$$

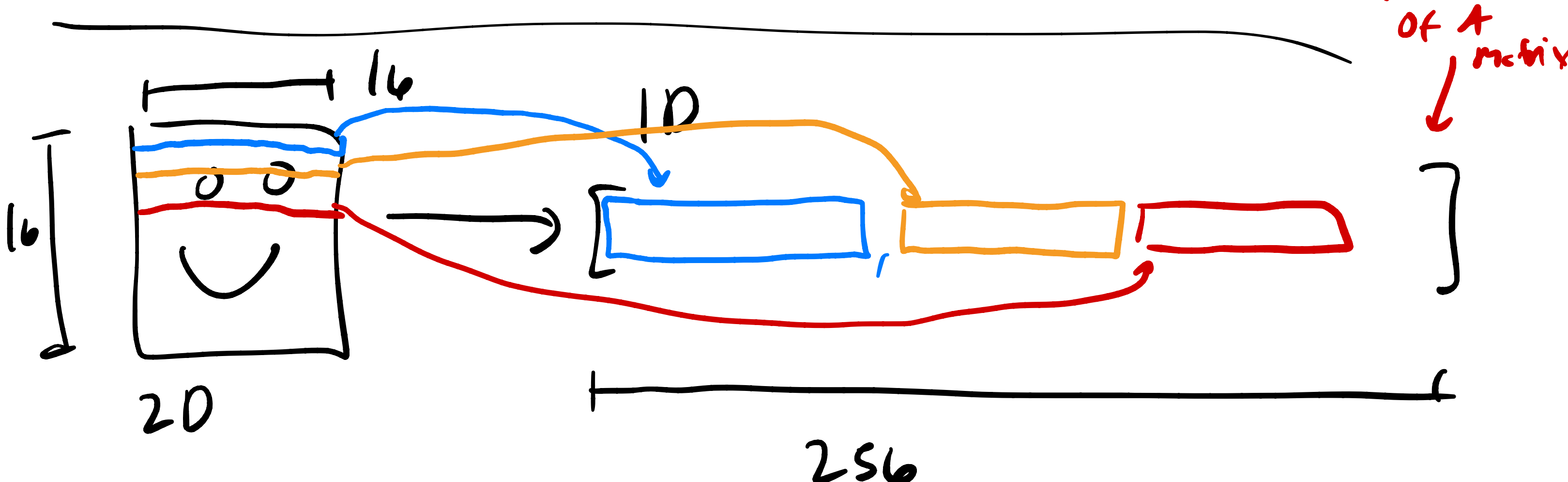
$\text{shape}(A_R) = \text{shape}(A)$

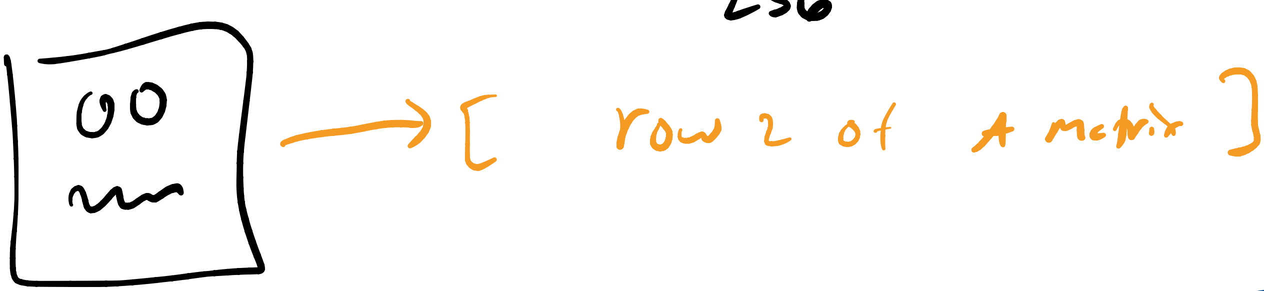
if we dropped PC2:



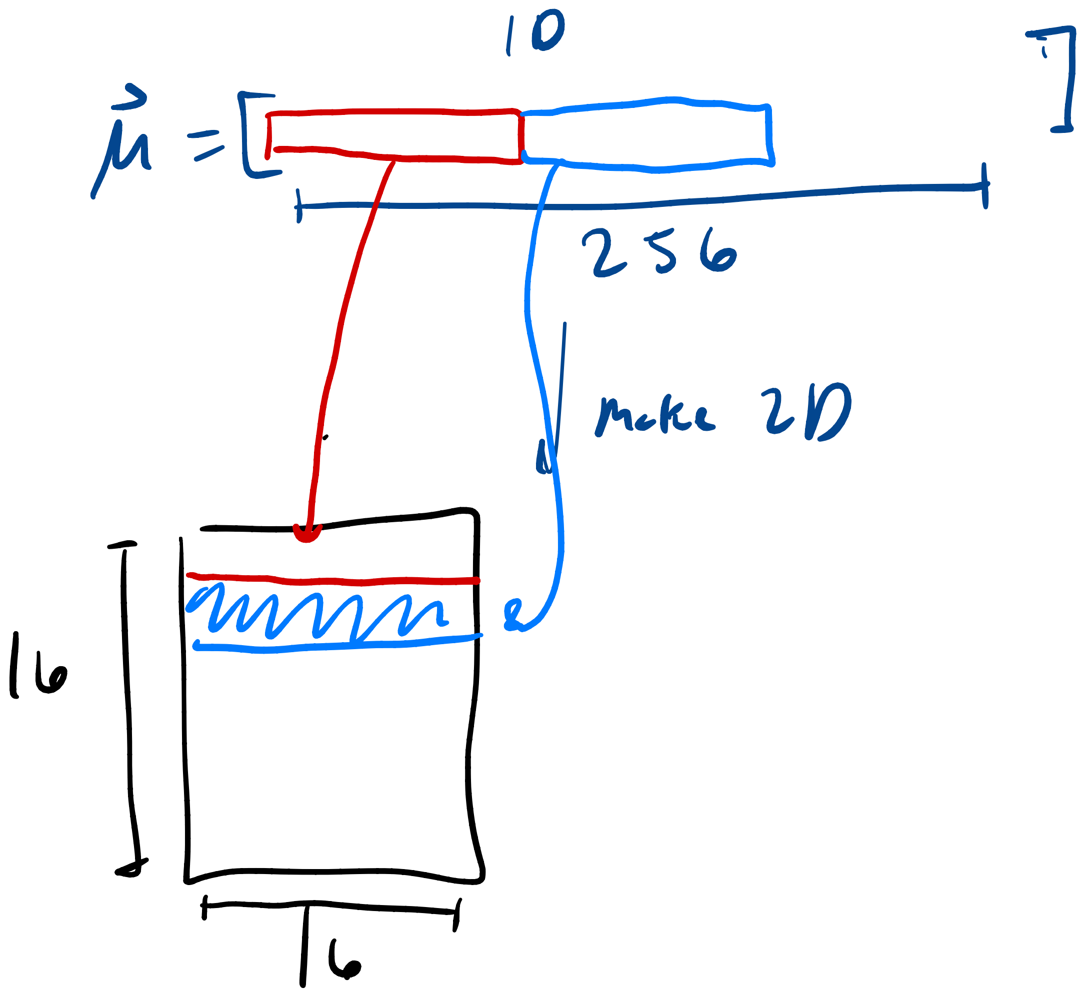
if we did <u>not</u> drop any PCs

$A_R = A$





$$\vec{\mu} = \underbrace{[\qquad\qquad\qquad]}_{256}$$



$\vec{\mu} =$
10
256
make 2D
16
16