Lecture 18

$$\underset{N\times q}{Z} = \underset{N\times p}{X}\ \underset{p\times q}{W}$$

$A = W^T X^T X W$

Find W to

(1) max total var

$$\text{TotVar}(Z) = \sum_{i=1}^{q} \text{Var}(z_i)$$ } max tr(A)

subject to

(2) $\text{Cor}(z_i, z_j) = 0$ for $i \neq j$ } A diag.

(3) $w_i = i^{th}$ col of W is a unit vec.

How do we do this?

$$\begin{aligned} A &= W^T X^T X W \\ &= W^T V D^T U^T U D V^T W \\ &= W^T V D^T D V^T W \\ &= W^T V \Delta^2 V^T W \end{aligned}$$

if $X = UDV^T$

$$D = \left[\begin{array}{c|c} D_* & 0 \\ \hline 0 & 0 \end{array}\right]$$

$D_* = \text{diag}(\sigma_1, \dots, \sigma_r)$

$$\Delta^2 = D^T D = \left[\begin{array}{c|c} D_*^2 & 0 \\ \hline 0 & 0 \end{array}\right] = \text{diag}(\sigma_1^2, \dots, \sigma_r^2, 0, \dots, 0)$$

Idea:

If $W = V$ then

$$= V^T V \Delta^2 V^T V$$

$$= \Delta^2 = \text{diag}$$

then $\mathrm{tr}(A) = \mathrm{tr}(\Delta^2) = \sum_{i=1}^{r} \sigma_i^2$ ← tot. var

Diagonalization is basically unique up to permutation of cols of W

Ex. $W = \begin{bmatrix} | & | & | & \\ v_2 & v_5 & v_3 & \cdots \\ | & | & | & \end{bmatrix}$ permutation

then instead $W^T V = V^T V = I$

I get $W^T V$ a permutation of I

$$W^T V = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & \\ 0 & 0 & 0 & 0 & 1 & \cdots \\ 0 & 0 & 1 & & \cdots & \end{bmatrix} = \Pi$$

So

$$A = \Pi \Delta^2 \Pi^T \leftarrow \text{still diagonal}$$

$$= \begin{bmatrix} \sigma_2^2 & & & \\ & \sigma_5^2 & & \\ & & \sigma_3^2 & \\ & & & \ddots \end{bmatrix}$$

but $\mathrm{tr}(A) = \sum_{i=1}^{n} \sigma_i^2$.

Problem: W is $p \times q$ but V is $p \times p$

I only want q cols in W

Soln: Choose cols of W to be same subset of cols of V

Claim: $W = \begin{bmatrix} | & | & | \\ v_2 & v_5 & v_3 \\ | & | & | \end{bmatrix}$ $q = 3$; $p = 5$

$$A = \underbrace{W^T V}_{\Pi} \Delta^2 V^T W$$

$$\Pi = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_2^2 & & \\ & \sigma_5^2 & \\ & & \sigma_3^2 \end{bmatrix}$$

Want to choose cols of V to max

$$tr(A) = \sum_{\text{cols chosen}} \sigma_j^2$$

Choose first q cols of V to be W

b/c then

$$tr(A) = \sum_{j=1}^{q} \sigma_j^2 .$$

Punchline: PCA

1. Typically, center cols of X (sometimes divide cols by SD)

2. $X = UDV^T$

(3) $Z = XW$

where $W = V_q \leftarrow$ first q cols of V

i.e. $z_i = X w_i$

$= X v_i$

$= \sigma_i u_i$

$$\left[\begin{array}{l} X = UDV^T \\ Xv_i = \sigma_i u_i \end{array}\right]$$

notice: $Var(z_i) = Var(\sigma_i u_i)$

$= \sigma_i^2 \underbrace{Var(u_i)}_{\propto 1}$

$\propto \sigma_i^2$

believe that u_i is mean centered

u_i and $mean(u_i) = 0$

$Var(u_i) = \frac{1}{N-1} u_i^T u_i$

$Var(u_i) \propto u_i^T u_i = 1$

in particular

$Var(z_i) = \frac{\sigma_i^2}{N-1}$

$Tot\ Var(Z) = \sum_{j=1}^{q} Var(z_i) = \frac{1}{N-1} \sum_{j=1}^{q} \sigma_j^2$

Pct. of orig. Var $= \dfrac{\frac{1}{N-1} \sum_{j=1}^{q} \sigma_j^2}{\frac{1}{N-1} \sum_{j=1}^{p} \sigma_j^2}$

$\leftarrow$ Var of PCs

var of orig. data

Comments:

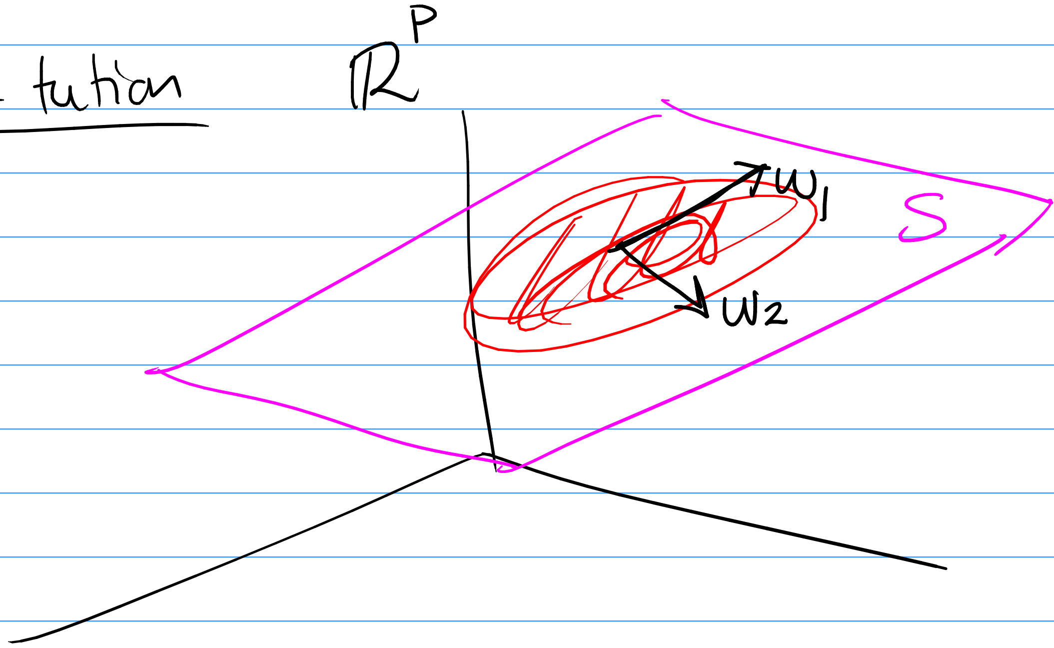
① $z_i \propto u_i$ in particular $z_i = \sigma_i u_i$

Kinda optional

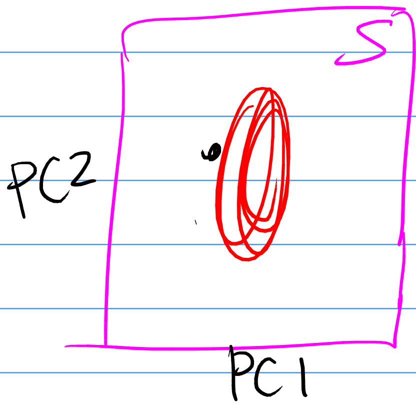
② mean centering X is also kinda optional

If I don't often u_1 = mean of X_is

Interpretation



w_i form the basis of this subspace in $\mathbb{R}^P$ and the z_i are the coordinate vectors of data in this new basis



Gen for a mtx A

$$P_A = A(A^TA)^{-1}A^{-1}$$

proj. onto Col A

if A has orthog. cols then

$$P_A = AA^T$$

pts proj. onto col W

$$X_q = XP_W \quad W = V_q$$

$$= \underbrace{XV_q}_{\text{coords i.e. } z}\underbrace{V_q^T}_{\text{basis}}$$

$$= UDV^TV_q \cdot V_q^T \qquad V^TV_q = \begin{bmatrix} I \\ \hline 0 \end{bmatrix}$$

V_q = first q cols of V

D_q = ''

U_q = ''

$$\underset{N \times P}{X_q} = U_q D_q V_q^T$$

truncated SVD

Theorem: Eckart-Young theorem

X_q is the best rank $= q$ approx. of X

$$X_q = \underset{B:\,\mathrm{rank}(B)=q}{\mathrm{argmin}} \|X - B\|$$