HW6

I PCA

Problem 1

Part 1 Show that $\sigma^2 I = \sigma^2 \sum_{i=1}^{p} \tilde{v}_i \tilde{v}_i^T$

i.e. show that $\sum_{i=1}^{p} \tilde{v}_i \tilde{v}_i^T = I$

Since $v \in R^p$, $\tilde{v}_i \in R^p$ $i \in [p]$

$$\sum_{i=1}^{p} \tilde{v}_i \tilde{v}_i^T = \sum_{i=1}^{p} \left[\begin{pmatrix} \tilde{v}_{i1} \\ \tilde{v}_{i2} \\ \vdots \\ \tilde{v}_{ip} \end{pmatrix} (\tilde{v}_{i1} \quad \tilde{v}_{i2} \quad \cdots \quad \tilde{v}_{ip}) \right]$$

$$= \sum_{i=1}^{p} \begin{bmatrix} \tilde{v}_{i1}^2 & \tilde{v}_{i1}\tilde{v}_{i2} & \cdots & \tilde{v}_{i1}\tilde{v}_{ip} \\ \tilde{v}_{i2}\tilde{v}_{i1} & \tilde{v}_{i2}^2 & \cdots & \tilde{v}_{i2}\tilde{v}_{ip} \\ \vdots & & \ddots & \vdots \\ \tilde{v}_{ip}\tilde{v}_{i1} & \tilde{v}_{ip}\tilde{v}_{i2} & \cdots & \tilde{v}_{ip}^2 \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^{p} \tilde{v}_{i1}^2 & \sum_{i=1}^{p} \tilde{v}_{i1}\tilde{v}_{i2} & \cdots & \sum_{i=1}^{p} \tilde{v}_{i1}\tilde{v}_{ip} \\ \sum_{i=1}^{p} \tilde{v}_{i2}\tilde{v}_{i1} & \sum_{i=1}^{p} \tilde{v}_{i2}^2 & \cdots & \sum_{i=1}^{p} \tilde{v}_{i2}\tilde{v}_{ip} \\ \vdots & & \ddots & \vdots \\ \sum_{i=1}^{p} \tilde{v}_{ip}\tilde{v}_{i1} & \sum_{i=1}^{p} \tilde{v}_{ip}\tilde{v}_{i2} & \cdots & \sum_{i=1}^{p} \tilde{v}_{ip}^2 \end{bmatrix}$$

$$= \begin{bmatrix} \langle \tilde{v}_1, \tilde{v}_1 \rangle & \langle \tilde{v}_1, \tilde{v}_2 \rangle & \cdots & \langle \tilde{v}_1, \tilde{v}_p \rangle \\ \langle \tilde{v}_2, \tilde{v}_1 \rangle & \langle \tilde{v}_2, \tilde{v}_2 \rangle & \cdots & \langle \tilde{v}_2, \tilde{v}_p \rangle \\ \vdots & & & \\ \langle \tilde{v}_p, \tilde{v}_1 \rangle & \langle \tilde{v}_p, \tilde{v}_2 \rangle & \cdots & \langle \tilde{v}_p, \tilde{v}_p \rangle \end{bmatrix}$$

Since $\tilde{v}_i$ are all orthogonal unit vectors $i \in [p]$

$$\begin{cases} \langle \tilde{v}_i, \tilde{v}_i \rangle = 1 \\ \langle \tilde{v}_i, \tilde{v}_j \rangle = 0 \quad (i \neq j) \end{cases}$$

then $\sum_{i=1}^{p} \tilde{v}_i \tilde{v}_i^T = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}_{p \times p} = I_p$

Thus $\sigma^2 I = \sigma^2 \sum_{i=1}^{p} \tilde{v}_i \tilde{v}_i^T$

Part 2 Show that $vv^T + \sigma^2 I = \|v\|_2^2 \tilde{v}_1 \tilde{v}_1^T + \sigma^2 \sum_{i=1}^{p} \tilde{v}_i \tilde{v}_i^T$

i.e. Show that $vv^T = \|v\|_2^2 \tilde{v}_1 \tilde{v}_1^T$

Since $\tilde{v}_1 = \frac{v}{\|v\|_2}$

Then $\|v\|_2^2 \tilde{v}_1 \tilde{v}_1^T = \|v\|_2^2 \cdot \frac{v}{\|v\|_2} \cdot \frac{v^T}{\|v\|_2} = \|v\|_2^2 \frac{vv^T}{\|v\|_2^2} = vv^T$

From Part 1 we show that $\sigma^2 I = \sigma^2 \sum_{i=1}^{p} \tilde{v}_i \tilde{v}_i^T$

Thus $vv^T + \sigma^2 I = \|v\|_2^2 \tilde{v}_1 \tilde{v}_1^T + \sigma^2 \sum_{i=1}^{p} \tilde{v}_i \tilde{v}_i^T$

Part 3 Show that $\Sigma = (\|v\|_2^2 + \sigma^2) \tilde{v}_1 \tilde{v}_1^T + \sigma^2 \sum_{i=2}^{p} \tilde{v}_i \tilde{v}_i^T$

$$\begin{aligned}
\Sigma &= vv^T + \sigma^2 I \\
&= \|v\|_2^2 \tilde{v}_1 \tilde{v}_1^T + \sigma^2 \sum_{i=1}^{p} \tilde{v}_i \tilde{v}_i^T \quad \text{(Part 2)} \\
&= \|v\|_2^2 \tilde{v}_1 \tilde{v}_1^T + \sigma^2 \tilde{v}_1 \tilde{v}_1^T + \sigma^2 \sum_{i=2}^{p} \tilde{v}_i \tilde{v}_i^T \\
&= (\|v\|_2^2 + \sigma^2) \tilde{v}_1 \tilde{v}_1^T + \sigma^2 \sum_{i=2}^{p} \tilde{v}_i \tilde{v}_i^T \\
&= \begin{pmatrix} | & | & & | \\ \tilde{v}_1 & \tilde{v}_2 & \cdots & \tilde{v}_p \\ | & | & & | \end{pmatrix}_{p \times p} \begin{pmatrix} \|v\|_2^2+\sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{pmatrix}_{p \times p} \begin{pmatrix} — & \tilde{v}_1^T & — \\ — & \tilde{v}_2^T & — \\ & \vdots & \\ — & \tilde{v}_p^T & — \end{pmatrix}_{p \times p}
\end{aligned}$$

Thus, the singular value decomposition of Σ is:

$\Sigma = VDV^T$

where $V \in R^{p \times p}$ is a square orthogonal matrix, $V^TV = VV^T = I$

with i^{th} column is singular vector $\tilde{v}_i$

$D \in R^{p \times p}$ is a square diagonal matrix with positive entries

$$\begin{cases} D_{11} = \|v\|_2^2 + \sigma^2 & \text{is the largest singular value} \\ D_{ii} = \sigma^2 \quad (i \neq 1) & \text{are other singular values} \end{cases}$$

Problem 2 Show that top left singular vector is $\frac{v}{\|v\|_2}$

In Problem 1 we show that the largest singular value of Σ is $\|v\|_2^2 + \sigma^2$

That is the first entry of diagonal matrix D : D_{11}

This singular value is corresponding to the first singular vector in orthogonal matrix V

That is $\tilde{v}_1 = \frac{v}{\|v\|_2}$

Problem 3

$x_i = v y_i + w_i \qquad y_i \sim N(0,1) \quad w_i \sim N(0, \sigma^2 I) \qquad x_i \sim N(0, v + \sigma^2 I) \in R^p$

$v \in R^p$

The covariance matrix of x_i : $Cov(x_i) \in R^{p \times p}$

$$Cov(x_i) = \begin{bmatrix} Cov(x_{i(1)}, x_{i(1)}) & Cov(x_{i(1)}, x_{i(2)}) & \cdots & Cov(x_{i(1)}, x_{i(p)}) \\ Cov(x_{i(2)}, x_{i(1)}) & Cov(x_{i(2)}, x_{i(2)}) & \cdots & Cov(x_{i(2)}, x_{i(p)}) \\ \vdots & & \ddots & \vdots \\ Cov(x_{i(p)}, x_{i(1)}) & Cov(x_{i(p)}, x_{i(2)}) & & Cov(x_{i(p)}, x_{i(p)}) \end{bmatrix}_{p \times p}$$

where $\begin{cases} \text{diagonal entries} \quad Cov(x_{i(j)}, x_{i(j)}) = Var(x_{i(j)}) \\ \text{non-diagonal entries} \quad Cov(x_{i(j)}, x_{i(k)}) \quad (j \neq k) \end{cases}$

Then compute these 2 kinds of entries:

① $Var(x_{i(j)}) = E[(x_{i(j)})^2] - E^2(x_{i(j)})$

$= E[(v_{(j)} y_i + w_{i(j)})^2] - 0$

$= E(v_{(j)}^2 y_i^2 + w_{i(j)}^2 + 2 v_{(j)} y_i w_{i(j)})$

$= v_{(j)}^2 E(y_i^2) + E(w_{i(j)}^2) + 2 v_{(j)} E(y_i w_{i(j)})$

$= v_{(j)}^2 [Var(y_i) + E^2(y_i)] + [Var(w_{i(j)}) + E^2(w_{i(j)})] + 2 v_{(j)} E(y_i) E(w_{i(j)})$

$= v_{(j)}^2 (1+0) + (\sigma^2 + 0) + 2 v_{(j)} \cdot 0 \cdot 0$ since y_i and w_i are uncorrelated

$= v_{(j)}^2 + \sigma^2$

② $Cov(x_{i(j)}, x_{i(k)})$

$$= E\left[\left[x_{i(j)} - E(x_{i(j)})\right]\left[x_{i(k)} - E(x_{i(k)})\right]\right]$$

$$= E\left[(x_{i(j)} - 0)(x_{i(k)} - 0)\right]$$

$$= E\left(x_{i(j)} x_{i(k)}\right)$$

$$= E\left[\left[v_{(j)} y_i + w_{i(j)}\right]\left[v_{(k)} y_i + w_{i(k)}\right]\right]$$

$$= E\left[v_{(j)} v_{(k)} y_i^2 + v_{(j)} y_i w_{i(k)} + w_{i(j)} v_{(k)} y_i + w_{i(j)} w_{i(k)}\right]$$

$$= v_{(j)} v_{(k)} E(y_i^2) + v_{(j)} E(y_i w_{i(k)}) + v_{(k)} E(w_{i(j)} y_i) + E\left(w_{i(j)} w_{i(k)}\right)$$

$$= v_{(j)} v_{(k)} + 0 + 0 + 0$$

$$= v_{(j)} v_{(k)}$$

Therefore

$$Cov(x_i) = \begin{bmatrix} v_{(1)}^2 + \sigma^2 & v_{(1)} v_{(2)} & \cdots & v_{(1)} v_{(p)} \\ v_{(2)} v_{(1)} & v_{(2)}^2 + \sigma^2 & \cdots & v_{(2)} v_{(p)} \\ \vdots & & \ddots & \vdots \\ v_{(p)} v_{(1)} & v_{(p)} v_{(2)} & \cdots & v_{(p)}^2 + \sigma^2 \end{bmatrix}_{p \times p}$$

$$= \begin{bmatrix} v_{(1)}^2 & v_{(1)} v_{(2)} & \cdots & v_{(1)} v_{(p)} \\ v_{(2)} v_{(1)} & v_{(2)}^2 & \cdots & v_{(2)} v_{(p)} \\ \vdots & & \ddots & \vdots \\ v_{(p)} v_{(1)} & v_{(p)} v_{(2)} & \cdots & v_{(p)}^2 \end{bmatrix}_{p \times p} + \sigma^2 \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & \cdots & 1 \end{bmatrix}_{p \times p}$$

$$= v v^T + \sigma^2 I_p \quad \text{for } v \in R^p$$

$$= \Sigma$$

The covariance of x_i is Σ.

Interpretation: because the PCA finds the most interesting direction of data, in this direction will keep the largest variance of data.

Since the covariance matrix of data is constructed by vector v

Then the projection on v will keep the most variance of data,

that is unit vector $\frac{v}{\|v\|_2}$ to be the top singular vector.

---

2 SVD

Problem 4

Part 1 Show ①② is equivalent

① The first way to produce u'

$$\begin{cases} u = \frac{Xv}{\|Xv\|_2} \\ v' = \frac{X^T u}{\|X^T u\|_2} \\ u' = \frac{Xv'}{\|Xv'\|_2} \end{cases}$$

plug in v'

$$\Rightarrow u' = \frac{X\frac{X^T u}{\|X^T u\|_2}}{\left\|X\frac{X^T u}{\|X^T u\|_2}\right\|_2} = \frac{\frac{XX^T u}{\|X^T u\|_2}}{\frac{\|XX^T u\|_2}{\|X^T u\|_2}} = \frac{XX^T u}{\|XX^T u\|_2}$$

plug in u

$$\Rightarrow u' = \frac{XX^T\frac{Xv}{\|Xv\|_2}}{\left\|XX^T\frac{Xv}{\|Xv\|_2}\right\|_2} = \frac{\frac{XX^TXv}{\|Xv\|_2}}{\frac{\|XX^TXv\|_2}{\|Xv\|_2}} = \boxed{\frac{XX^TXv}{\|XX^TXv\|_2}}$$

② The Second way to produce u'

$$\begin{cases} \tilde{u} = Xv \\ \tilde{v}' = X^T\tilde{u} \\ u' = \frac{X\tilde{v}'}{\|X\tilde{v}'\|_2} \end{cases}$$

plug in $\tilde{v}'$

$$\Rightarrow u' = \frac{XX^T\tilde{u}}{\|XX^T\tilde{u}\|_2}$$

Plug in $\tilde{u}$

$$\Rightarrow u' = \boxed{\frac{XX^TXv}{\|XX^TXv\|_2}}$$ the same as first way

Part 2 Show equalities of u and u'

$$\begin{cases} u = \dfrac{Xv}{\|Xv\|_2} \\ u' = \dfrac{X\tilde{v}'}{\|X\tilde{v}'\|_2} \end{cases}$$

$$X = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T \qquad \sigma_1 > \sigma_2 > 0$$

$$\begin{cases} \alpha_1 = v^T v_1 > 0 \\ \alpha_2 = v^T v_2 > 0 \end{cases}$$

① Compute u

$$u = \frac{Xv}{\|Xv\|_2} = \frac{(\sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T) v}{\|(\sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T) v\|_2} = \frac{\sigma_1 u_1 v_1^T v + \sigma_2 u_2 v_2^T v}{\|\sigma_1 u_1 v_1^T v + \sigma_2 u_2 v_2^T v\|_2}$$

since $\begin{cases} \alpha_1 = v^T v_1 = v_1^T v > 0 \\ \alpha_2 = v^T v_2 = v_2^T v > 0 \end{cases}$

$$u = \frac{\sigma_1 \alpha_1 u_1 + \sigma_2 \alpha_2 u_2}{\|\sigma_1 \alpha_1 u_1 + \sigma_2 \alpha_2 u_2\|_2}$$

$$= \frac{\sigma_1 \alpha_1 u_1 + \sigma_2 \alpha_2 u_2}{\sqrt{\sigma_1^2 \alpha_1^2 \sum_j^p u_{1(j)}^2 + \sigma_2^2 \alpha_2^2 \sum_j^p u_{2(j)}^2}}$$

$$= \frac{\sigma_1 \alpha_1 u_1 + \sigma_2 \alpha_2 u_2}{\sqrt{\sigma_1^2 \alpha_1^2 + \sigma_2^2 \alpha_2^2}} \qquad (u_1, u_2 \text{ are unit vectors})$$

$$\boxed{= \frac{u_1 + \dfrac{\alpha_2 \sigma_2}{\alpha_1 \sigma_1} u_2}{\sqrt{1 + \dfrac{\alpha_2^2 \sigma_2^2}{\alpha_1^2 \sigma_1^2}}}}$$

② Compute u'

$$\tilde{u} = Xv = \sigma_1 \alpha_1 u_1 + \sigma_2 \alpha_2 u_2$$

$$\tilde{v}' = X^T \tilde{u}$$

$$= (\sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T)^T (\sigma_1 \alpha_1 u_1 + \sigma_2 \alpha_2 u_2)$$

$$= (\sigma_1 v_1 u_1^T + \sigma_2 v_2 u_2^T)(\sigma_1 \alpha_1 u_1 + \sigma_2 \alpha_2 u_2)$$

$$= \sigma_1^2 \alpha_1 v_1 \underline{u_1^T u_1} + \sigma_1 \sigma_2 \alpha_2 v_1 \underline{u_1^T u_2} + \sigma_1 \sigma_2 \alpha_1 v_2 \underline{u_2^T u_1} + \sigma_2^2 \alpha_2 v_2 \underline{u_2^T u_2}$$

Since $u_1^T u_2 = 0$ and u_1, u_2 are unit vectors

$$\tilde{v}' = \sigma_1^2 \alpha_1 v_1 + \sigma_2^2 \alpha_2 v_2$$

$$u' = \frac{X \tilde{v}'}{\|X \tilde{v}\|_2}$$

First compute $X\tilde{v}'$

$$X\tilde{v}' = (\sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T)(\sigma_1^2 \alpha_1 v_1 + \sigma_2^2 \alpha_2 v_2)$$

$$= \sigma_1^3 \alpha_1 u_1 \underline{v_1^T v_1} + \sigma_1 \sigma_2^2 \alpha_2 u_1 \underline{v_1^T v_2} + \sigma_1^2 \sigma_2 \alpha_1 u_2 \underline{v_2^T v_1} + \sigma_2^3 \alpha_2 u_2 \underline{v_2^T v_2}$$

Since $v_1^T v_2 = 0$ and v_1, v_2 are unit vectors

$$X\tilde{v}' = \sigma_1^3 \alpha_1 u_1 + \sigma_2^3 \alpha_2 u_2$$

Then $u' = \dfrac{\sigma_1^3 \alpha_1 u_1 + \sigma_2^3 \alpha_2 u_2}{\|\sigma_1^3 \alpha_1 u_1 + \sigma_2^3 \alpha_2 u_2\|_2}$

$$= \frac{\sigma_1^3 \alpha_1 u_1 + \sigma_2^3 \alpha_2 u_2}{\sqrt{(\sigma_1^3 \alpha_1)^2 + (\sigma_2^3 \alpha_2)^2}}$$

$$= \frac{\sigma_1^3 \alpha_1 u_1 + \sigma_2^3 \alpha_2 u_2}{\sqrt{\sigma_1^6 \alpha_1^2 + \sigma_2^6 \alpha_2^2}}$$

$$\boxed{= \frac{u_1 + \dfrac{\alpha_2 \sigma_2^3}{\alpha_1 \sigma_1^3} u_2}{\sqrt{1 + \dfrac{\alpha_2^2 \sigma_2^6}{\alpha_1^2 \sigma_1^6}}}}$$

Part 3 Show that $[u']^T u_1 > u^T u_1$

① left side $= [u']^T u_1$

plug in u' computed in Part 2

$$\text{left side} = \left[\frac{u_1 + \frac{\alpha_2 \sigma_2^3}{\alpha_1 \sigma_1^3} u_2}{\sqrt{1 + \frac{\alpha_2^2 \sigma_2^6}{\alpha_1^2 \sigma_1^6}}} \right]^T u_1$$

$$= \frac{u_1^T + \frac{\alpha_2 \sigma_2^3}{\alpha_1 \sigma_1^3} u_2^T}{\sqrt{1 + \frac{\alpha_2^2 \sigma_2^6}{\alpha_1^2 \sigma_1^6}}} \cdot u_1$$

$$= \frac{1}{\sqrt{1 + \frac{\alpha_2^2 \sigma_2^6}{\alpha_1^2 \sigma_1^6}}} \left[u_1^T u_1 + \frac{\alpha_2 \sigma_2^3}{\alpha_1 \sigma_1^3} u_2^T u_1 \right]$$

Since u_1 is unit vector and $u_1^T u_2 = u_2^T u_1 = 0$

$$\text{left side} = \frac{1}{\sqrt{1 + \frac{\alpha_2^2 \sigma_2^6}{\alpha_1^2 \sigma_1^6}}}$$

② right side $= u^T u_1$

plug in u and u_1 computed in Part 2

$$\text{right side} = \left[\frac{u_1 + \frac{\alpha_2 \sigma_2}{\alpha_1 \sigma_1} u_2}{\sqrt{1 + \frac{\alpha_2^2 \sigma_2^2}{\alpha_1^2 \sigma_1^2}}} \right]^T \cdot u_1$$

$$= \frac{u_1^T + \frac{\alpha_2 \sigma_2}{\alpha_1 \sigma_1} u_2^T}{\sqrt{1 + \frac{\alpha_2^2 \sigma_2^2}{\alpha_1^2 \sigma_1^2}}} \cdot u_1$$

$$= \frac{1}{\sqrt{1 + \frac{\alpha_2^2 \sigma_2^2}{\alpha_1^2 \sigma_1^2}}} \left(u_1^T u_1 + \frac{\alpha_2 \sigma_2}{\alpha_1 \sigma_1} u_2^T u_1 \right)$$

Since u_1 is unit vector and $u_1^T u_2 = u_2^T u_1 = 0$

$$\text{right side} = \frac{1}{\sqrt{1+\frac{\alpha_2^2\sigma_2^2}{\alpha_1^2\sigma_1^2}}}$$

$$\frac{\text{left side}}{\text{right side}} = \frac{1}{\sqrt{1+\frac{\alpha_2^2\sigma_2^6}{\alpha_1^2\sigma_1^6}}} \div \frac{1}{\sqrt{1+\frac{\alpha_2^2\sigma_2^2}{\alpha_1^2\sigma_1^2}}}$$

$$= \frac{\sqrt{1+\frac{\alpha_2^2\sigma_2^2}{\alpha_1^2\sigma_1^2}}}{\sqrt{1+\frac{\alpha_2^2\sigma_2^6}{\alpha_1^2\sigma_1^6}}}$$

$$= \frac{\sqrt{1+\frac{\alpha_2^2}{\alpha_1^2}\left(\frac{\sigma_2}{\sigma_1}\right)^2}}{\sqrt{1+\frac{\alpha_2^2}{\alpha_1^2}\left(\frac{\sigma_2}{\sigma_1}\right)^6}}$$

since $\sigma_1 > \sigma_2 > 0$

then $0 < \frac{\sigma_2}{\sigma_1} < 1$, $\left(\frac{\sigma_2}{\sigma_1}\right)^2 > \left(\frac{\sigma_2}{\sigma_1}\right)^6 > 0$

then numerator $>$ denominator > 0

Then $\frac{\text{left side}}{\text{right side}} > 1$

Therefore $[u']^T u_1 > u^T u_1$

Extra credit

$$X = USV^T = \sum_{j=1}^{p \wedge n} u_j v_j^T \sigma_j \qquad X \in R^{p \times n}$$

$$S = U^T X V \qquad U \in R^{p \times p} \quad V \in R^{n \times n} \quad S \in R^{p \times n}$$

$$= \begin{pmatrix} — u_1^T — \\ — u_2^T — \\ \vdots \\ — u_p^T — \end{pmatrix} \times \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix}$$

$$= \begin{pmatrix} u_1^T X v_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & u_2^T X v_2 & \cdots & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & u_p^T X v_p & \cdots & 0 \end{pmatrix}_{p \times n}$$

$$= \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \sigma_2 & & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots & & \vdots \\ 0 & \cdots & & \sigma_p & \cdots & 0 \end{pmatrix}_{p \times n}$$

since the diagonal entry of rectangular diagonal matrix S : σ_j are the singular values in decreasing order

$$\sigma_j = u_j^T X v_j$$

where u_j is the jth column vector of orthogonal matrix U

v_j is the jth column vector of orthogonal matrix V

Then $\sigma_1 = \max_{\|u\|_2 \le 1, \, \|v\|_2 \le 1} u^T X v$