

MATH 312
LECTURE 17

Singular Value Decomposition

From last lecture,

Spectral Theorem: M real and symmetric \Leftrightarrow

$$M = QDQ^T, \text{ with } D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

$$Q = \left[\begin{array}{c|c|c} \vec{f}_1 & \dots & \vec{f}_n \end{array} \right] \text{ orthogonal matrix.}$$

\hookrightarrow orthonormal eigenvectors.

Example: Find $M = QDQ^T$ for $M = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$

Eigenvalues: $M - \lambda I = \begin{bmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{bmatrix} \rightsquigarrow$

$$\rightsquigarrow (2-\lambda)^3 + 1 + 1 - (2-\lambda) - (2-\lambda) - (2-\lambda) = 0 \Leftrightarrow$$

$$\Leftrightarrow (4 + \lambda^2 - 4\lambda)(2-\lambda) + 2 - 6 + 3\lambda = 0 \Leftrightarrow$$

$$\Leftrightarrow 8 - 4\lambda + 2\lambda^2 - \lambda^3 - 2\lambda + 4\lambda^2 - 4 + 3\lambda = 0 \Leftrightarrow$$

$$\Leftrightarrow -\lambda^3 + 6\lambda^2 - 9\lambda + 4 = 0 \sim \lambda = 1$$

$$-\lambda^3 + 6\lambda^2 - 9\lambda + 4 \quad \begin{array}{l} \lambda-1 \\ \hline -\lambda^2 + 5\lambda - 4 \end{array}$$

$$\begin{array}{r} \lambda^3 - \lambda^2 \\ \hline 5\lambda^2 - 9\lambda + 4 \\ -5\lambda^2 + 5\lambda \\ \hline -4\lambda + 4 \\ 4\lambda - 4 \\ \hline 0 \end{array}$$

$$-\lambda^3 + 6\lambda^2 - 9\lambda + 4 = (\lambda-1)(-\lambda^2 + 5\lambda - 4)$$

$$-\lambda^2 + 5\lambda - 4 = 0$$

$$\Leftrightarrow \lambda = \frac{-5 \pm \sqrt{25 - 16}}{-2} = \frac{-5 \pm 3}{-2} \begin{array}{l} \nearrow 4 \\ \searrow 1 \end{array}$$

$$\Sigma \lambda_1 = 1 = \lambda_2, \lambda_3 = 4.$$

Eigenvektoren:

$$\underline{\lambda=1}: A - J = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \rightarrow \vec{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

Bed $\vec{v}_1, \vec{v}_2 \neq 0!$

→ Gram-Schmidt:

$$\vec{u}_1 = \vec{v}_1$$

$$\vec{u}_2 = \vec{v}_2 - \left(\frac{\vec{v}_2 \cdot \vec{u}_1}{\|\vec{u}_1\|^2} \right) \vec{u}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ -1 \end{bmatrix} \quad (\text{CGK } \vec{u}_1 \cdot \vec{u}_2 = 0)$$

∅

$$\vec{y}_1 = \frac{\vec{u}_1}{\|\vec{u}_1\|} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}, \quad \vec{y}_2 = \frac{\vec{u}_2}{\|\vec{u}_2\|} = \frac{1}{\sqrt{\frac{1}{2} + 1}} \begin{bmatrix} 1/2 \\ 1/2 \\ -1 \end{bmatrix}$$

Finally,

$$K - 4J = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \rightarrow \begin{bmatrix} -2 & 1 & 1 \\ 0 & -3/2 & 3/2 \\ 0 & 3/2 & -3/2 \end{bmatrix} \rightarrow \begin{bmatrix} -2 & 1 & 1 \\ 0 & -3 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

↑
free

$$\begin{cases} \text{So } x_3 = \alpha \\ x_2 = x_3 = \alpha \\ -2x_1 = -\alpha - \alpha \rightarrow x_1 = \alpha \end{cases} \left\{ \begin{array}{l} \vec{u}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \sim \vec{f}_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \end{array} \right.$$

(Check that

$$\text{So } Q = \left[\vec{f}_1 \mid \vec{f}_2 \mid \vec{f}_3 \right], \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}, \quad (\vec{f}_1 \cdot \vec{f}_3 = 0 = \vec{f}_2 \cdot \vec{f}_3).$$

• Why is this theorem also called Principal Axis Theorem?

We can identify symmetric matrices with quadratic forms, that is, (homogeneous) polynomials of the form

$$ax^2 + 2bxy + cy^2 = 1 \quad (\text{in 2d})$$

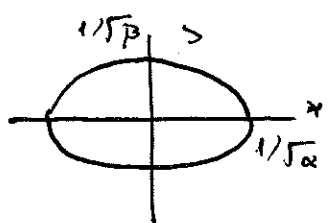
$$\text{Since } ax^2 + 2bxy + cy^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

So given a 2×2 matrix, we can think of the corresponding polynomial $\vec{x}^T M \vec{x}$.

Example: Draw the ellipse $5x^2 + 8xy + 5y^2 = 1$.

Sol: This is not easy right now.

We know how to draw ellipses given by $\alpha x^2 + \beta y^2 = 1$



$$\left(\frac{x}{1/\sqrt{\alpha}}\right)^2 + \left(\frac{y}{1/\sqrt{\beta}}\right)^2 = 1$$

But notice that

$$5x^2 + 8xy + 5y^2 = 1 = \begin{bmatrix} x & y \end{bmatrix} \underbrace{\begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}}_{M \text{ symmetric}} \begin{bmatrix} x \\ y \end{bmatrix}$$

So we can write $M = Q D Q^T$. In this particular case,

$$Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \rightarrow \vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$D = \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix}$$

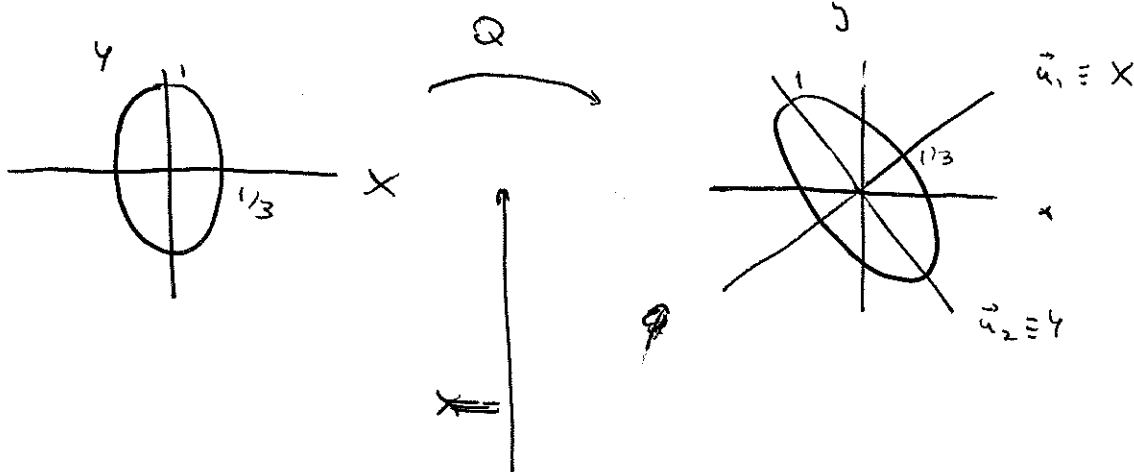
That is, the ellipse is given by

$$\vec{x}^T \Lambda \vec{x} = \vec{x}^T Q \Lambda Q^T \vec{x} = (Q^T \vec{x})^T \Lambda (Q^T \vec{x})$$

If we change variables $\begin{bmatrix} x \\ y \end{bmatrix} = \vec{X} = Q^T \vec{x} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \rightsquigarrow$

$$\rightsquigarrow \begin{cases} X = \frac{x+y}{\sqrt{2}} \\ Y = \frac{x-y}{\sqrt{2}} \end{cases} \left\{ \text{ellipse} = 9X^2 + Y^2 = 1, \right.$$

so in X-Y axis:



This is a linear (a orthogonal) transformation.

$$\begin{bmatrix} 1/3 \\ 0 \end{bmatrix} \rightsquigarrow Q \begin{bmatrix} 1/3 \\ 0 \end{bmatrix} \rightsquigarrow \begin{bmatrix} \vec{u}_1 & | & \vec{u}_2 \end{bmatrix} \begin{bmatrix} 1/3 \\ 0 \end{bmatrix} = \frac{1}{3} \vec{u}_1 //$$

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \rightsquigarrow Q \begin{bmatrix} 0 \\ 1 \end{bmatrix} \rightsquigarrow \dots = \vec{u}_2 //$$

* Linear transformation given by Q orthogonal:

→ Preserves lengths: $\|Q\vec{x}\|^2 = (Q\vec{x})^T (Q\vec{x}) = \vec{x}^T Q^T Q \vec{x} = \vec{x}^T \vec{x} = \|\vec{x}\|^2$

→ Preserves angles: $(Q\vec{x}) \cdot (Q\vec{y}) = \vec{x}^T Q^T Q \vec{y} = \vec{x}^T \vec{y} = \vec{x} \cdot \vec{y} \Rightarrow \cos(\widehat{\vec{x}, \vec{y}}) = \cos(\widehat{Q\vec{x}, Q\vec{y}})$

VII Positive semidefinite matrices

This is a preparation to SVD.

Def: Positive semidefinite matrix.

~~Matrix~~ An $n \times n$ symmetric matrix is called positive semidefinite if

$$\vec{x}^T M \vec{x} \geq 0 \quad \forall \vec{x} \in \mathbb{R}^n.$$

Def: It is called positive definite if $\vec{x}^T M \vec{x} > 0 \quad \forall \vec{x} \in \mathbb{R}^n, \vec{x} \neq \vec{0}$.

Proposition: M is posit. semidef. \Leftrightarrow all λ 's ≥ 0 .

Proof.

1) \Rightarrow Let λ be eig: $M \vec{u} = \lambda \vec{u} \Rightarrow \vec{u}^T M \vec{u} = \lambda \vec{u}^T \vec{u} \Rightarrow$
 $\Rightarrow \lambda \|\vec{u}\|^2 = \vec{u}^T M \vec{u} \geq 0 \Rightarrow \lambda \geq 0.$

2) \Leftarrow Let all $\lambda \geq 0$.

$$M \text{ symmetric} \Rightarrow M = Q D Q^T$$

Any \vec{x} can be written in the basis $\{\vec{z}_1, \dots, \vec{z}_n\}$.

$$\vec{x} = \alpha_1 \vec{z}_1 + \dots + \alpha_n \vec{z}_n. \text{ Then,}$$

$$\begin{aligned}
\vec{x}^T M \vec{x} &= (\alpha_1 \vec{g}_1^T + \dots + \alpha_n \vec{g}_n^T) (\lambda_1 \vec{g}_1 \vec{g}_1^T + \dots + \lambda_n \vec{g}_n \vec{g}_n^T) (\alpha_1 \vec{g}_1 + \dots + \alpha_n \vec{g}_n) = \\
&= (\alpha_1 \vec{g}_1^T + \dots + \alpha_n \vec{g}_n^T) (\lambda_1 \alpha_1 \vec{g}_1 + \lambda_2 \alpha_2 \vec{g}_2 + \dots + \lambda_n \alpha_n \vec{g}_n) = \\
&= \lambda_1 \alpha_1^2 + \dots + \lambda_n \alpha_n^2 \geq 0 \\
&\quad \uparrow \\
&\quad \lambda_1, \dots, \lambda_n \geq 0 // .
\end{aligned}$$

Exercise: ~~The matrices~~ Let A be an $m \times n$ matrix with $r = \text{rank}(A)$. Show that

1) AA^T and $A^T A$ have the same nonzero eigenvalues.

2) $\text{rank}(AA^T) = \text{rank}(A^T A) = r$.

3) Both are positive semidefinite.

Sol:

1) Let λ, \vec{z} eig. of $A^T A$ ($\lambda \neq 0$): $A^T A \vec{z} = \lambda \vec{z}$

Then,

$$AA^T A \vec{z} = \lambda A \vec{z} \rightsquigarrow (AA^T)(A \vec{z}) = \lambda (A \vec{z}).$$

So, if $A \vec{z} \neq 0$, then $A \vec{z}$ is eigenvector of AA^T with eigenvalue λ .

↳ But if $A \vec{z} = 0$, then $\lambda = 0$ (and we said $\lambda \neq 0$).

We can do the same starting from AA^T .

2) $\text{rank}(AA^T) = \text{rank}(A^T A)$ from 1) and 3).

Now, let's show $\text{rank}(A^T A) = r = \text{rank}(A)$.

By rank-nullity theorem, since $A^T A$ is $n \times n$ as A , we just need to show that $\dim(N(A^T A)) = \dim N(A)$.

But $N(A^T A) = N(A)$!

$$\left\{ \begin{array}{l} \rightarrow \vec{x} \in N(A) \Rightarrow A\vec{x} = \vec{0} \Rightarrow A^T A\vec{x} = \vec{0} \Rightarrow \vec{x} \in N(A^T A). \\ \rightarrow \vec{x} \in N(A^T A) \Rightarrow A^T A\vec{x} = \vec{0} \Rightarrow A\vec{x} \in N(A^T) = C(A)^\perp. \end{array} \right.$$

$$\text{But } A\vec{x} \in C(A) \rightarrow \text{So } \overset{\checkmark}{A\vec{x} = \vec{0}} \Rightarrow \vec{x} \in N(A).$$

$$3) \vec{x}^T A^T A \vec{x} = (A\vec{x})^T A\vec{x} = \|A\vec{x}\|^2 \geq 0 \quad \forall \vec{x}$$

$$\vec{x}^T A A^T \vec{x} = (A^T \vec{x})^T A^T \vec{x} = \|A^T \vec{x}\|^2 \geq 0 \quad \forall \vec{x}.$$

Chapter 7: Singular Value Decomposition.

We will see that this decomposition has plenty of applications.

All of them share one idea: the information contained in a big matrix is usually reduced, as many columns are "almost" linearly dependent.

Think of this matrix:

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} [1 \ 1 \ 1 \ 1] \rightarrow \text{rank}(A) = 1.$$

and \swarrow the rows are "almost" lin. dependent

$$\begin{aligned} B &= \begin{bmatrix} 1.1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 1.1 & 1.2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 0.1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.1 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0.2 & 0 & 0 \end{bmatrix} = \\ &= \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} [1 \ 1 \ 1 \ 1] + \begin{bmatrix} 0.1 \\ 0 \\ 0.1 \end{bmatrix} [1 \ 0 \ 0 \ 0] + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} [0 \ 0.2 \ 0 \ 0] \\ &\quad \underbrace{\hspace{10em}}_{\text{are "negligible"}} \end{aligned}$$

B has $\text{rank}(B) = 3$, but we would like a method that tells us that $\text{rank}(B) \approx 1$ (that is, that detected the "main" part of B).

So, our goal will be to find how to decompose a matrix as a sum of rank one matrices, in a way that we can, which are the important pieces.

Goal: Given A $m \times n$, find

$$A = U \Sigma V^T = \sigma_1 \vec{u}_1 \vec{v}_1^T + \dots + \sigma_r \vec{u}_r \vec{v}_r^T,$$

where U is orthogonal $m \times m$ matrix.

V is orthogonal $n \times n$ matrix.

Σ is $m \times n$ diagonal matrix with $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq 0$.

$$U \begin{bmatrix} \vec{u}_1 & \dots & \vec{u}_m \end{bmatrix} \quad \Sigma \begin{bmatrix} \sigma_1 & & \\ & \sigma_r & \\ & & \dots \end{bmatrix} \quad V \begin{bmatrix} \vec{v}_1 & \dots & \vec{v}_n \end{bmatrix}$$

I'll first say how to find U, Σ, V and then we will show why:

2) V

Find orthonormal eigenvectors for $A^T A$:

→ r from $\lambda_1, \dots, \lambda_r \rightsquigarrow \vec{v}_1, \dots, \vec{v}_r$ (Don't forget to make them
with size
→ $n-r$ from $\lambda=0 \rightsquigarrow \vec{v}_{r+1}, \dots, \vec{v}_n$ (You might need
Gram-Schmidt).

$$V = \left[\begin{array}{c|c} \vec{v}_1 & \dots & \vec{v}_r & \vec{v}_{r+1} & \dots & \vec{v}_n \end{array} \right] \text{ (orthogonal) }_{n \times n}$$

3) U

For $i=1, \dots, r$, let $\vec{u}_i = \frac{1}{\sigma_i} A \vec{v}_i$

Then, find $n-r$ orthonormal eigenvectors corresponding to $\lambda=0$. That is, find an orthonormal basis for $N(AA^T) (=N(A^T))$.

$$U = \left[\begin{array}{c|c|c|c} \frac{A \vec{v}_1}{\sigma_1} & \dots & \frac{A \vec{v}_r}{\sigma_r} & \vec{u}_{r+1} & \dots & \vec{u}_n \end{array} \right] \text{ (orthogonal) }_{m \times m}$$

↖

$$(A = U \Sigma V^T \Rightarrow AV = \Sigma U)$$

To complete the previous example, u, v ?

• $V \rightarrow$ eigenvectors of $A^T A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ ($\lambda_1 = 3, \lambda_2 = 1$)

$$A^T A - 3I = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \rightarrow \vec{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rightarrow \vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ (unit size)}$$

$$A^T A - I = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \sim \vec{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \rightarrow \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ (unit size)}$$

$$\text{So } V = [\vec{v}_1 \ \vec{v}_2] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

• u : it is a 3×3 matrix, and $r = \text{rank}(A) = 2 \rightarrow$

$$\vec{u}_1 = \frac{1}{\sigma_1} A \vec{v}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \frac{1}{\sqrt{2}} = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \rightarrow \text{(this are already unit size, we'll see why)}$$

$$\vec{u}_2 = \frac{1}{\sigma_2} A \vec{v}_2 = \frac{1}{1} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \frac{1}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Finally, \vec{u}_3 is an eigenvector of AA^T corresponding to $\lambda = 0$:

$$AA^T = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{matrix} x_3 = \alpha \\ x_2 = -\alpha \\ x_1 = \alpha \end{matrix}$$

↑
free

$$\text{So } \vec{u}_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

In summary, $\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$, $U = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{3} \\ 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix}$,

$$V = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

One can check $A = U\Sigma V^T$; columns of U orthogonal;
columns of V orthogonal \checkmark