EECS 127 . Lecture 4 . Jan 30, 2020.

Today:
- Properties of Symmetric matrices.
- Finish PCA
- SVD.

Admin
- HW1 due.
- Concurrent Enrollment.

Spectral Theorem for Symmetric matrices.

$A \in S^n$, λ_i are the eigenvalues.

(1) $\lambda_i \in \mathbb{R}$

(2) Eigenspaces corresponding to distinct eigenvalues are orthogonal.

(3) Algebraic multiplicity of λ_i = Geometric multiplicity.

Symm. Mat. are diagonalizable.

$$A = U \Sigma U^T$$

Σ = diagonal matrix.

U, U^T; are orthornormal.

Variational characterization of Eigenvalues. (4.3.)

$A \in S^n$. $\qquad \dfrac{\vec{x}^T A \vec{x}}{\vec{x}^T \vec{x}}$ "Rayleigh coef"

Thm: $\lambda_{min}(A) \le \dfrac{\vec{x} A \vec{x}}{\vec{x}^T \vec{x}} \le \lambda_{max}(A) \quad \forall \vec{x} \in \mathbb{R}^n \neq 0.$

- $\lambda_{max}(A) = \max_{\|\vec{x}\|_2 = 1} \vec{x}^T A \vec{x}$
- $\lambda_{min}(A) = \min_{\|\vec{x}\| = 1} \vec{x} A \vec{x}$

Proof: $A = U \Lambda U^T$

$$\vec{x}^T A \vec{x} = \vec{x}^T U \Lambda U^T \vec{x}$$

$$= \vec{y}^T \Lambda \vec{y}$$

$$= \sum_{i=1}^{n} \lambda_i y_i^2$$

$\vec{y} = U^T \vec{x} \qquad \vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$

$[y_1\, y_2] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$

$\lambda_{min} = \lambda_{min} \sum y_i^2 \le$ — " —— $\le \lambda_{max} \cdot \sum y_i^2 = \lambda_{max}$

∎

- **Positive Semidefinite Matrix**

$A \succeq 0$ PSD

$A \succ 0$ PD

$A \in S^n$ is said to be PSD.

if $\vec{x}^T A \vec{x} \geq 0$ for all $\vec{x} \in \mathbb{R}^n$.

- **Matrix Square Root**

$A \succeq 0$ ⇔ $A = B^T B$

$B \succeq 0$ symmetric

$A = U \Lambda U^T$

$B = U \Lambda^{1/2} U^T$

⇔ $A = \underbrace{(U \Lambda^{1/2})}_{= B^T} \underbrace{\Lambda^{1/2} U^T}_{B}$

Back to PCA

We showed that our problem was equivalent to.

$$\vec{w}^* = \arg\max_{\|\vec{w}\|_2 = 1} \vec{w}^T C \vec{w}$$

C = Covariance Matrix of our dat.

$C = X^T X$.

$$X = \begin{bmatrix} - \vec{x}_1^T - \\ - \vec{x}_2^T - \\ \vdots \\ - \vec{x}_n^T - \end{bmatrix}$$

λ_{max}, $\vec{w}^*$ = corresponding eigenvector.

Singular Value Decomposition.

$A \in \mathbb{R}^{m \times n}$ Rank r.

$$A = \sigma_1 \vec{u}_1 \vec{v}_1^T + \sigma_2 \vec{u}_2 \vec{v}_2^T + \cdots + \sigma_r \vec{u}_r \vec{v}_r^T$$

Convention: $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$

$\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r \in \mathbb{R}^m$ and are orthonormal.

$\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r \in \mathbb{R}^n$ — " ——.

$$A = U \Sigma V^T$$ "Compact" SVD.

$$= \underbrace{\begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_r \end{bmatrix}}_{m \times r} \underbrace{\begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{bmatrix}}_{r \times r} \underbrace{\begin{bmatrix} — \vec{v}_1^T — \\ — \vec{v}_2^T — \\ \vdots \end{bmatrix}}_{r \times n}$$

$$= U_{m \times m} \underbrace{\begin{bmatrix} \Sigma_{r \times r} & 0 \\ 0 & 0 \end{bmatrix}}_{m \times n} V^T_{n \times n}$$

• How to find SVD?

SVD of A, look at eigenvalues of A^TA.

$A^TA \in \mathbb{R}^{n \times n}$ Symmetric.

① Real evals. ② e-vectors are $\perp$.

$B = A^TA$

$B^T = (A^TA)^T = A^TA$

$\vec{x}^T A^T A \vec{x} \geq 0$

$\|A\vec{x}\|^2 \geq 0$

① $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_r > 0$ eigenvalues of A^TA

② $\vec{v}_1, \vec{v}_2 \cdots \vec{v}_r$ are corresponding e-vectors. Orthonormal.

$$A^TA\vec{v}_i = \lambda_i \vec{v}_i$$

③ $\sigma_i = \sqrt{\lambda_i}$, Define $\vec{u}_i$

$$A\vec{v}_i = \sigma_i \vec{u}_i \qquad i = 1, 2, \cdots r.$$

Claim: $\vec{u}_i$ are orthonormal.

$$(\sigma_i \vec{u}_i)^T (\sigma_j \vec{u}_j) = (A\vec{v}_i)^T (A\vec{v}_j) = \vec{v}_i^T A^TA\vec{v}_j = \vec{v}_i^T \lambda_j \vec{v}_j$$

$$\sigma_i \sigma_j \vec{u}_i^T \vec{u}_j = \vec{v}_i^T \lambda_j \vec{v}_j$$

$i = j$: $\sqrt{\lambda_i}\sqrt{\lambda_i}\, \|\vec{u}_i\|^2 = \lambda_i \|\vec{v}_i\|^2 \Rightarrow \|\vec{u}_i\|^2 = 1$

$i \neq j$: $\vec{u}_i^T \vec{u}_j = 0$

(4)

$$A \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_r \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & u_r \\ | & | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_r \end{bmatrix}$$

$$\underset{m\times n}{A} \; \underset{n\times r}{V_r} = \underset{m\times r}{U_r} \; \underset{r\times r}{\Sigma}$$

$$A = A \, V_r V_r^T = U_r \Sigma V_r^T$$

$V_r : n\times r \to n\times n$

$$\underset{n\times n}{V} = [V_r \;\; V_1]$$

Consider $VV^T = [V_r \;\; V_1] \begin{bmatrix} V_r^T \\ V_1^T \end{bmatrix} = V_r V_r^T + V_1 V_1^T$

$$A \cdot VV^T = A V_r V_r^T + A V_1 V_1^T$$

$$A = A V_r V_r^T$$

Consider $A V_1 V_1^T$.

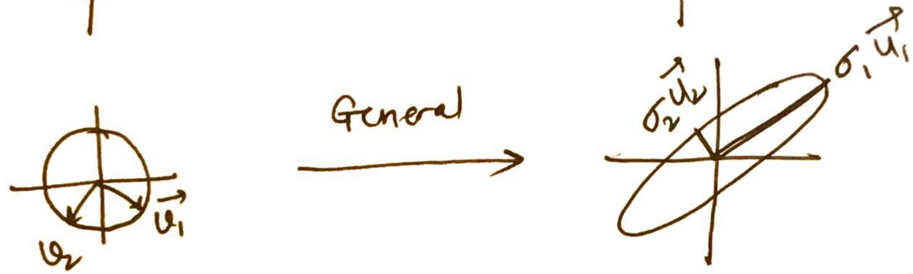
$A^T A V_1 = 0$

$V_1^T A^T A V_1 = 0.$

$(AV_1)^T (AV_1) = 0.$

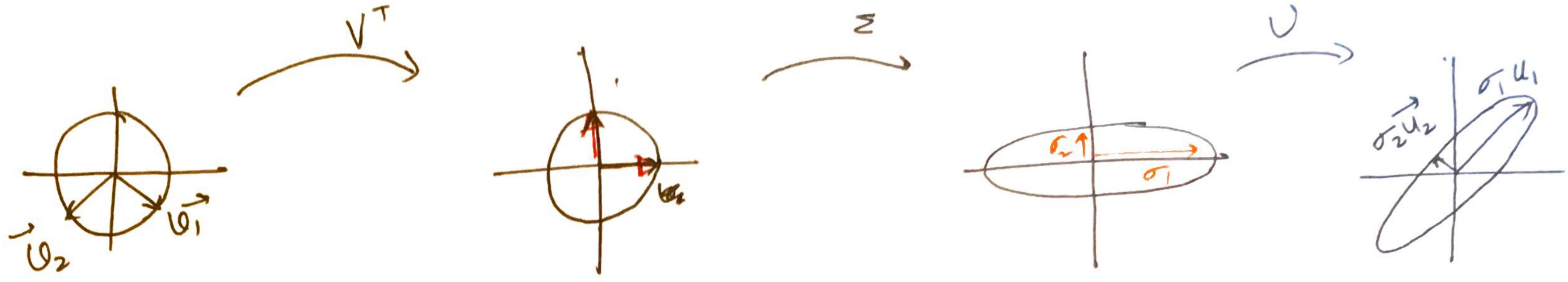
- Null(A) = Null(ATA)

$A = U \Sigma V^T$



If you start with e-vectors of $A^T A$

$$\underset{\|\vec{x}\|=1}{\operatorname{argmax}} \|A\vec{x}\| = \underset{\|\vec{x}\|=1}{\operatorname{argmax}} \vec{x}^T A^T A \vec{x}$$



$\vec{v}_1$
$\vec{v}_2$
V^T
Σ
σ_2
σ_1
U
$\sigma_1 u_1$
$\sigma_2 \vec{u}_2$