10. SVD – SINGULAR VALUE DECOMPOSITION

- It is one of the most important decompositions and it is guaranteed to exist to any matrix.
- Has some similarities with the Eigen Value Decomp. (EVD); however, recall that EVD only applies to a class of matrices (Diagonalizable Matrices);
- Has several applications

10.1. Conceptual construction

it is built in terms of the concepts of singular values σ and the corresponding singular vectors u and v.

$\sigma \in \mathbb{R}^+$ is a singular value ~~for~~ for unit vectors u, v iff $(\|u\| = \|v\| = 1)$

$$\left.\begin{array}{l} (1)\ Au = \sigma v \\ (2)\ A^* v = \sigma u \end{array}\right\}$$

in this case
u is the r-svec of A, σ
v is the l-svec of

(σ, u) is a right singular pair } for A.
(σ, v) is a left singular pair }

Now, pre-multiply eqs (1) & (2) by A^* and A respectively.

$$A^* A u = \sigma (A^* v) = \sigma \overbrace{(\sigma u)}^{(2)} = \sigma^2 u$$

$$\boxed{A^* A u = \sigma^2 u}$$
: ~~the r-sing~~ the r-sing. pair (σ, u) for (A) is the ~~th~~ eigen pair (σ^2, u) for $A^* A$

$$\boxed{\begin{array}{l} \lambda(A^*A) = \lambda(AA^*) = \sigma^2(A) \\ \text{or } \sigma(A) = \lambda^{1/2}(A^*A) = \lambda^{1/2}(AA^*) \end{array}}$$

$$A^* v = \sigma u$$

$$AA^* v = \sigma \underbrace{Au}_{(1)}$$

$$AA^* v = \sigma(\sigma v)$$

$$\boxed{AA^* v = \sigma^2 v}$$
: the l-sing pair (σ, v) for A is ~~the~~ eigen pair (σ^2, v) for AA^*

therefore, we can calculate the singular values (3)
$\sigma(A)$, the l-svec and the r-svec from a standard ~~eval~~ eigen problem in terms of matrices A^*A and AA^* (they are different in general)

1) It is easy to show that $\lambda(A^*A) = \lambda(AA^*)$

2) Right s-vec u in $Au = \sigma v$ is obtained from the eigenvec problem

$$A^*A u = \sigma^2 u$$

3) Left s-vec v in $A^* v = \sigma u$ is obtained from

$$AA^* v = \sigma^2 v$$

Matrices A^*A and AA^* have some nice properties (better working with $\left\{\begin{matrix} A^*A \\ AA^* \end{matrix}\right\}$ than A)

a) $\left.\begin{matrix} (A^*A)^* = A^*(A^*)^* = A^*A \\ (AA^*)^* = (A^*)^* A^* = AA^* \end{matrix}\right\}$ they are hermitian $\therefore \lambda(A^*A) = \lambda(AA^*) \in \mathbb{R}$

b) they are positive semi-definite: $A^*A \geq 0$, $AA^* \geq 0$

From the definition $B = B^*$ is positive (semi) def. iff

$x^* B x \geq 0 \iff \lambda(B) \geq 0$ (evals of a pos semdfi B are nonneg.)

$$x^*(A^*A)x = x^*A^*Ax = (Ax)^*(Ax) = y^*y = \|y\|^2 \geq 0$$

$$x^*(AA^*)x = x^*AA^*x = (A^*x)^*(A^*x) = z^*z = \|z\|^2 \geq 0$$

c)

c) they are normal matrices : a normal matrix F is such that

$$F^*F = FF^* \quad (\text{square})$$

and it is guaranteed to have a full set of eigenvectors. A^*A and AA^* are different normal matrices.

a) $A^*A \triangleq B$: $B^*B = BB^*$ (Normality Condition)

but $(A^*A)^* = A^*A$ (hermitian)

then $(A^*A)^* \{A^*A\} = A^*A(A^*A)^*$ $\therefore$ A^*A is a (square) normal matrix

b) $AA^* \triangleq C$: $C^*C = CC^*$

$(AA^*)^* \{AA^*\} = AA^*(AA^*)$ $\therefore$ AA^* is another (square) normal matrix

if $A_{N\times M}$, $\left.\begin{array}{l} A^*A \text{ is } M\times M \\ AA^* \text{ is } N\times N \end{array}\right\}$ $\therefore$ $A^*A \neq AA^*$ in general

A.

A^*A provide a full set of orthonormal evecs which form an unitary matrix of right sing. vectors for A

AA^* provide a full set of ortho. evecs which form an unitary matrix of left sing. vectors for A.

10.2. the fundamental theorem

$r = \mathrm{rank}(A)$

thm: Let $A \in \mathbb{C}_r^{M\times N}$. Then there exist unitary (orthogonal) matrices $U \in \mathbb{C}^{M\times M}$ and $V \in \mathbb{C}^{N\times N}$ such that

$$A = U \Sigma V^*$$

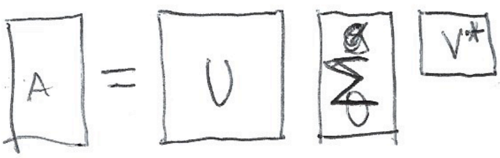
where $\Sigma = \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}$ and $S = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 \cdots & \\ 0 & & \sigma_r \end{bmatrix}$

$\sigma_1 \geq \sigma_2 \geq \sigma_3 \cdots \geq \sigma_r > 0$ (this is a convention. Order is arbitrary in the theorem)

Specifically

$$A_{M\times N} = \underbrace{\begin{bmatrix} \underbrace{U_1}_{r} & \underbrace{U_2}_{M-r} \end{bmatrix}}_{M \times M} \begin{array}{c} \\ r \\ M-r \end{array}\begin{bmatrix} \overset{r}{S} & \overset{N-r}{0} \\ 0 & 0 \end{bmatrix} \underbrace{\begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix}}_{N}\begin{array}{c} r \\ N-r \end{array} = U_1 S V_1^*$$

compact/thin SVD



Submatrices are all determined by ~~the~~ $r = \mathrm{rank}(A)$

$U_1 \in \mathbb{C}^{M\times r}$ $\qquad$ $V_1 \in \mathbb{C}^{N\times r}$

$U_2 \in \mathbb{C}^{M\times(M-r)}$ $\qquad$ $V_2 \in \mathbb{C}^{N\times(N-r)}$

I delivered in class the definitions (on-the-fly) of

- ~~Normal matrices~~: A is Normal if $A^*A = AA^*$

A normal matrix ~~is~~ has a full set of eigenvectors. As such, it is diagonalizable

- Hermitian Matrices: $A^* = A$, $A^T = A$ (symmetric)

Hermitian matrices have real eigenvalues

$Ax = \lambda x \quad \cancel{\Rightarrow x^* A x = }$

$x^* A x = \lambda \|x\|^2 \quad (1)$

$(x^* A x)^* = \bar{\lambda} \|x\|^2$

$x^* A^* x = \bar{\lambda} \|x\|^2$

~~(1) = (2)~~ but $A = A^*$

then $x^* A^* x = x^* A x = \bar{\lambda} \|x\|^2$

$x^* A x = \bar{\lambda} \|x\|^2 \quad (2)$

$(1) - (2): \quad 0 = (\lambda - \bar{\lambda}) \|x\|^2$

since $x \neq 0 \Leftrightarrow \|x\|^2 \neq 0$

$\lambda - \bar{\lambda} = 0 \Leftrightarrow \boxed{\lambda = \bar{\lambda}}$

- Positive ~~Def. Ma~~ (semi) Definite matrices

A Hermitian matrix (symmetric) $A^* = A$ is positive (semi) def. and denoted $A \geq 0$ if $\lambda(A) \geq 0 \Leftrightarrow x^* A x \geq 0, x \neq 0$

Likewise, $A \leq 0$ is Negative (semi) definite if $\lambda(A) \leq 0 \Leftrightarrow x^* A x \leq 0, x \neq 0$

A indefinite matrix has $\lambda(\cdot)$ both $\lambda \geq 0$ and $\lambda \leq 0$

$A \geq 0: \quad x^* A x = x^* U \Lambda U^* x = z^* z, \; z = \Lambda^{1/2} U^* x, \; = \|z\|^2 \geq 0$

$A \leq 0: \quad x^* A x = y^* \bar{\Lambda} y, \; y = U^* x; \; \bar{\Lambda} \leq 0 \therefore \bar{\Lambda} = -\Lambda \geq 0$

$y^* \bar{\Lambda} y = -y^* \Lambda y \leq 0$ since $y^* \Lambda y \geq 0$.

Proof: Start with A^*A (same holds for AA^*) ⑥

$A^*A v_i = \sigma_i^2 v_i \quad i=1,r$

$A^*A [v_1\ v_2 \cdots v_r] = [v_1 \cdots v_r] \begin{bmatrix} \begin{bmatrix}\sigma_1^2 & & \\ & \ddots & \\ & & \sigma_r^2\end{bmatrix} & \\ & 0 \end{bmatrix}$

$A^*A [V_1\ \cancel{V_2}] = [V_1\ \cancel{V_2}] \begin{bmatrix} S^2 & \cancel{0} \\ \cancel{0} & \cancel{0} \end{bmatrix}$

$A^*A V_1 = V_1 S^2$ (pre-mult. by V_1^*)

$V_1^* A^* A V_1 = \underbrace{V_1^* V_1}_{I_r} S^2$

σ_i^2 are eigenvalues of A^*A.
rank-r matrix A^*A: $\sigma_1^2 \geq \sigma_2^2 \geq \cdots \sigma_r^2 > 0$
~~...~~ r positive σ_i^2, $M-r$ zeros for $M>N$

$V_1^* A^* A V_1 = S^2$ pre & pos mult S^{-1}

$$\boxed{S^{-1} V_1^* A^* A V_1 S^{-1} = I_r}$$

Now $A^*[A V_2] = [V_2 \cdot \underbrace{0}]$ pre $\times V_2^*$

$\sigma = 0$ zero sing vals

$V_2^* A^* A V_2 = 0$

missing step $\|Av_i\|^2 = 0$

$(AV_2)^* AV_2 = 0 \quad \therefore \|AV_2\|^2 = 0$ (matrix Norms)

$\Rightarrow AV_2 = 0$

$(AV_2)^*(AV_2) = B^*B$ this is not a matrix Norm. $Tr(B^*B)$ is!

Now: $A = U\Sigma V^*$ $(\times V)$

$AV = U\Sigma = A[V_1\ V_2] = [U_1\ U_2] \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}$

$[AV_1\ \ AV_2] = [U_1 S\ \ 0]$ or $\begin{cases} AV_2 = 0 \text{ (shown above)} \\ AV_1 = U_1 S \end{cases}$

thus $AV_1 = U_1 S$ $(\times S^{-1})$

~~$AV_2 = 0$~~

$AV_1 S^{-1} = U_1 \Rightarrow$ Gusto $U_1 = AV_1 S^{-1}$.

Correct, but might be tautological.

Now choose $U_1 \triangleq A V_1 S^{-1} \in \mathbb{C}^{M \times r}$

then, from before:

$$\underbrace{S^{-1} V_1^* A^*}_{U_1^*} \underbrace{A V_1 S^{-1}}_{U_1} = I_r \Bigg\} \quad U_1^* U_1 = I_r$$

ortho cols, but U_1 is Not unitary

Now complete U_1 so that

$U = [U_1 \ U_2]$ is unitary

(check LAUB 96 or random vecs + G-S)

then

$$U^* A V = \begin{bmatrix} U_1^* A V_1 & U_1^* A V_2 \\ U_2^* A V_1 & U_2^* A V_2 \end{bmatrix} = \begin{bmatrix} U_1^* A V_1 & 0 \\ U_2^* A V_1 & 0 \end{bmatrix}$$

because $AV_2 = 0$

we know $U_1 = A V_1 S^{-1}$

$U_1^* U_1 = U_1^* A V_1 S^{-1}$

$I_r = U_1^* A V_1 S^{-1}$

$I_r \cdot S = U_1^* A V_1 S^{-1} S$

$\therefore \boxed{U_1^* A V_1 = S}$

Now: $U_2^* A V_1 = U_2^* (U_1 S) = \underbrace{U_2^* U_1}_{} S = 0 \cdot S = 0$

cols of $U = [U_1 \ U_2]$ are orthonormal

$\therefore U_2^* A V_1 = 0$

FINALLY

$$U^* A V = \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix} = \Sigma$$

$$U^* A V = \Sigma \iff A = U \Sigma V^*$$

■

10.3: Some Applications

1) Outer product expansion

$A = \sum_{i=1}^{r} \sigma_i u_i v_i^*$. the truncated SVD outer expansion is the best $\| \ \|_2$ ~~approxi~~ k-rank approximation for A:

$$A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^* \qquad k < r$$

so that $\|A - A_k\|_2$ is minimum.

2) Pseudo-inverse via SVD

For any $A \in \mathbb{C}_r^{m \times n}$: $A = U \Sigma V^*$. then

$$A^+ = V \Sigma^+ U^* \quad , \quad \Sigma^+ = \begin{bmatrix} S^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

Proof: test Moore-Penrose Conditions (LAUB 30)

3) Bases for the fundamental spaces

a) U_1 is ortho basis for $R(A)$ $(R(U_1) = R(A))$

b) U_2 is $\perp$ basis for $R(A^*)$ $(R(U_2) = R(A)^{\perp})$

c) V_1 is ortho basis for $N(A)^{\perp}$ $(R(V_1) = N(A)^{\perp})$

d) V_2 is ortho basis for $N(A)$ $(R(V_2) = N(A))$

4) Lin sys sol

$Ax = b$

$A = U \Sigma V^*$ $\qquad U \Sigma V^* x = b$

$$\left.\begin{aligned} \Sigma V^* x &= U^* b \\ \Sigma y &= U^* b \end{aligned}\right\} \quad y = \underbrace{\Sigma^{-1}}_{\text{simple inverse}} U^* b \Rightarrow \left\{\begin{aligned} V^* x &= y \\ V V^* x &= V y \end{aligned}\right.$$

$\therefore \boxed{x = V y}$ it's performed only with unitary transf.

Eigenvalue problem: generic matrix $A \in \mathbb{C}^{N\times N}$

$Ax = \lambda x$: (λ, x) r-pair $\qquad y^*A = \mu y^*$: (μ, y^*) l-pair

$\lambda(A) = \{\lambda \mid |A-\lambda I| = 0\}$ $\qquad \mu(A) = \{\mu \mid |A-\mu I| = 0\}$

$Ax - \lambda x = 0 \quad (A-\lambda I)x = 0$ $\qquad y^*A = \mu y^* \quad y^*(A-\mu I) = 0$

$|A-\lambda I| = 0$ $\qquad |A-\mu I| = 0$

$\therefore \boxed{\lambda(A) = \mu(A)}$ left & right evals are equal for $A \in \mathbb{C}^{N\times N}$

Generic matrix $A \in \mathbb{C}^{N\times N}$

Different $\lambda_i \in \lambda(A)$ produce ortho left and right evecs

$\lambda_k \neq \lambda_e$: $y_e^* A = \lambda_e y_e^*$

$$A x_k = \lambda_k x_k$$

$$\underbrace{y_e^* A}_{} x_k = \lambda_k y_e^* x_k$$

$$\lambda_e y_e^* x_k = \lambda_k y_e^* x_k$$

$$\lambda_e y_e^* x_k - \lambda_k y_e^* x_k = 0$$

$$(\lambda_e - \lambda_k) y_e^* x_k = 0$$

since $\lambda_e \neq \lambda_k \therefore \boxed{y_e^* x_k = 0}$

Generic Matrix $A \in \mathbb{C}^{N\times N}$ that have a distinct set of $\{\lambda_k\}$ have N-LI revecs and N-LI levecs (they are LI, not necessarily <u>ortho</u>

Hermitian matrix $A = A^* \in \mathbb{C}^{N\times N}$

- Real evals: $\lambda(A) \in \mathbb{R}$ (Not necessarily positive)
- has a full set of evecs (because it is a normal matrix)
- their evecs are orthogonal to diff. $\lambda_i \neq \lambda_e$