Similar matrices

The discussion on similar matrices covers non-symmetric or symmetric matrices.

A_{nxn} and B_{nxn} are similar matrices if for some matrix M:

$$B = M^{-1}AM$$

Example: $S^{-1}AS = \Lambda$ so A is similar to Λ because the relation $S^{-1}AS = \Lambda$ holds. So, I can find B (a family of matrices) which are similar to A if $B = M^{-1}AM$. This means that Λ is a special member of that family (the best one). Because if $B = M^{-1}AM$ then I still can write $A = MBM^{-1}$ which looks exactly as the $S\Lambda S^{-1}$ diagonalization of the matrix A.

In other words, if B is Λ then M is S (the eigenvectors matrix).

Example: $A = \begin{matrix} 2 & 1 \\ 1 & 2 \end{matrix}$

we can find $\Lambda = \begin{matrix} 3 & 0 \\ 0 & 1 \end{matrix}$

A and Λ are similar.

Also,

$$\underbrace{\begin{bmatrix} 1 & -4 \\ 0 & 1 \end{bmatrix}}_{M^{-1}} \underbrace{\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}}_{\substack{A \\ \lambda = 3,\ 1}} \underbrace{\begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}}_{M} = \begin{bmatrix} 1 & -4 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 9 \\ 1 & 6 \end{bmatrix} = \underbrace{\begin{bmatrix} -2 & -15 \\ 1 & 6 \end{bmatrix}}_{\substack{B \\ \lambda = 3,\ 1}}$$

So, there is something that connects all matrices with $\lambda = 3, 1$. Similar matrices have same eigenvalues but possibly different eigenvectors. How can we find the connection between the eigenvectors of A and the eigenvectors of B?

$$\begin{aligned} A\,x &= \lambda x \\ A\ I\,x &= \lambda x \\ A\,MM^{-1}x &= \lambda x \\ M^{-1}AMM^{-1}x &= \lambda M^{-1}x \\ (M^{-1}AM)M^{-1}x &= \lambda M^{-1}x \\ B\ \underbrace{M^{-1}x}_{} &= \lambda \underbrace{M^{-1}x}_{} \\ B\ \text{vector} &= \lambda\ \text{vector} \end{aligned}$$

vector is obviously represents the eigenvectors of B. Therefore, If B and A are similar matrices (i.e. have the same eigenvalues) then eigenvectors of B = M^{-1} *(eigenvectors of A).

That is why Λ is a special member in a family of similar matrices – because its eigenvectors matrix is $I = \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}$

Example: is the matrix $B_1 = \begin{bmatrix} 3 & 7 \\ 0 & 1 \end{bmatrix}$ similar to the matrix A used above? What about the matrix $B_2 = \begin{bmatrix} 1 & 7 \\ 0 & 3 \end{bmatrix}$?

All we have to do is to find out the eigenvalues of B_1 and B_2.

<u>Special case</u>: when A is not diagonalizable. For example $A=\begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$

Compare the case with $B = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$

Not similar

Both A and B have $\lambda_1 = \lambda_2 = 4$ but A is not diagonalizable, then we can't say that A and B are similar matrices. Therefore, B is alone in its own family (B is similar to itself. In other words, B belongs to a small family). If A was diagonalizable then it would have been in the same family as B.

Singular value decomposition (Motivation)

We have seen what seemed to be defects in matrices.

1. When an nxn matrix, A, doesn't have a full rank (i.e. has free columns which also means: $1 \le \dim(\text{null}(A)) \le n$).

2. When a matrix is not diagonalizable in the form $S\Lambda S^{-1}$.

In fact the second defect is more problematic because inability to diagonalize a matrix will devote the matrix to very poor computational properties. While rank deficiency won't. Remember that the two defects stated above are independent. Meaning, you may be able to diagonalize a rank deficient matrix (think of a projection matrix). On the other hand, you may not be able to diagonalize a full rank matrix (think of this matrix $\begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$).

It is therefore crucial for efficient computation to be able to diagonalize any matrix. Diagonalization in the form $S\Lambda S^{-1}$ is conditioned (A must be square, and must have n independent eigenvectors). Singular value decomposition solves this limitations; we can diagonalize any matrix using SVD.

Singular Value Decomposition

Diagonalizing a matrix A to $A = S\Lambda S^{-1}$ is conditioned by having independent eigenvectors (such that S^{-1} exists). It also means that A is a square matrix (Because eigenvalues are only defined for square matrices).

$$A = S \Lambda S^{-1}$$

must have independent vectors (eigen-vectors of A)

diagonal matrix with eigenvalues on the diagonal

If A is symmetric then its eigenvectors are $\perp$ which means that S becomes Q

$$A = Q\Lambda Q^{-1}$$
$$= Q\Lambda Q^{T}$$

you can see the advantage that a symmetric matrix has; (no inverse needed as $Q^{-1} = Q^{T}$).

if A is a pos. def. (a special class of Symmetric matrices) then λ's are all positive

eigenvector

indep. | dep.

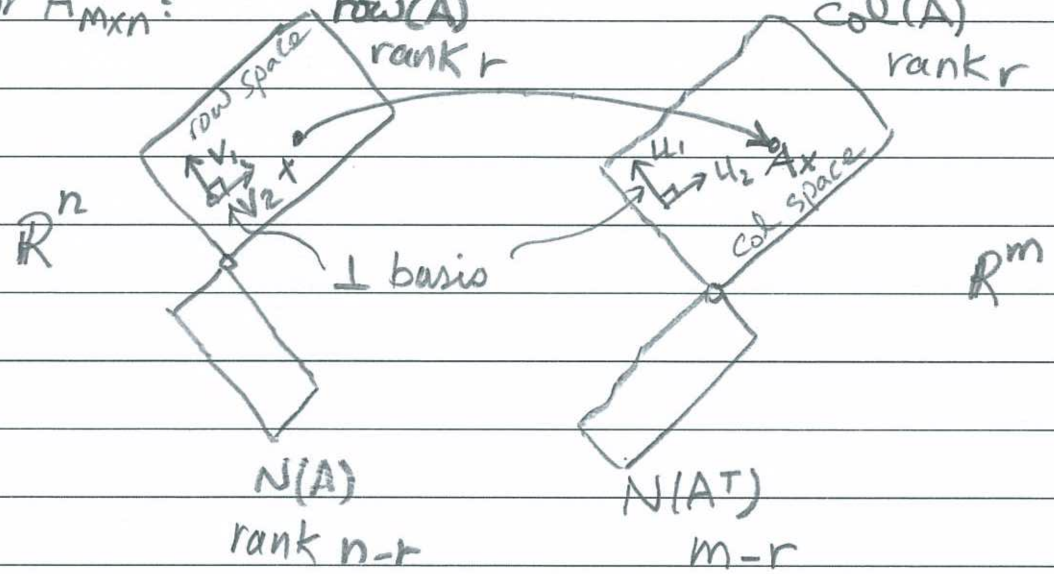
$A = Q\Lambda Q^{T}$ (A symmetric) | $A = S\Lambda S^{-1}$ (A non-symmetric)

We would like to diagonalize any matrix

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

U: orthonormal; Σ: diag.; V^T: orthonormal.

Let's remember the picture of linear Transformation for $A_{m\times n}$:



if V_1 is in rowspace(A) then $AV_1 = u_1$ where u_1 is going to be in Col(A)

Now if we choose V_1, V_2 as $\perp$ basis in row(A), then how can we find a way to take them to u_1, u_2 ($\perp$ basis in col(A)).

$$A V_1 = \sigma_1 u_1$$

$$A V_2 = \sigma_2 u_2$$

$$A \underbrace{[V_1 \; V_2 \; \cdots \; V_r]}_{\substack{\text{orthogonal} \\ \text{basis for} \\ \text{row}(A) \\ r \le n}} = \underbrace{[u_1 \; u_2 \; \cdots \; u_r]}_{\substack{\text{orthogonal} \\ \text{basis for} \\ \text{Col}(A) \\ r \le n}} \begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_2 & \\ 0 & & \ddots \; \sigma_r \end{bmatrix}$$

So the goal now is to find $AV = U\Sigma$

$$A = U\Sigma V^{-1}$$

since V is $\perp$ then $V^{-1} = V^T$:

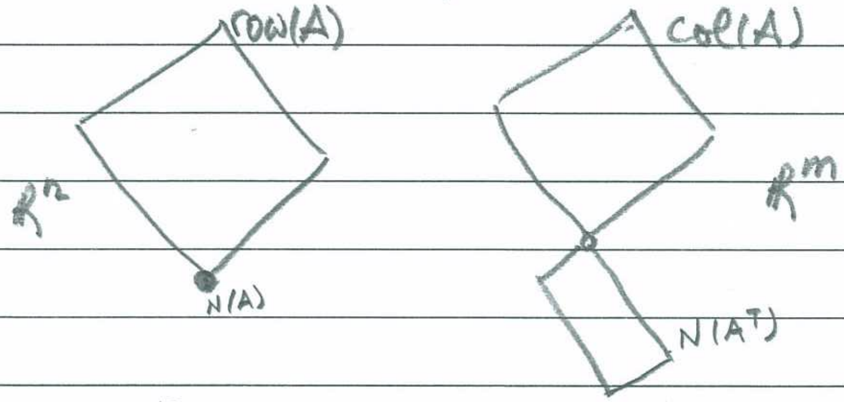
$$A = U \Sigma V^T$$

in $\mathbb{R}^m$
left singular vectors

diagonal singular values

in $\mathbb{R}^n$
right singular vectors.

Let's study the case when $A_{m\times n}$ and
$\operatorname{rank}(A) = r = n < m$ $\quad (N(A) = 0)$ (i.e. no null space)



$$A\,[v_1 \cdots v_r] = [u_1 \cdots u_r\ u_{r+1} \cdots u_m] \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{bmatrix}$$

$$\underset{m\times n}{A} = \underset{m\times m}{U}\ \underset{m\times n}{\Sigma}\ \underset{n\times n}{V^T}$$

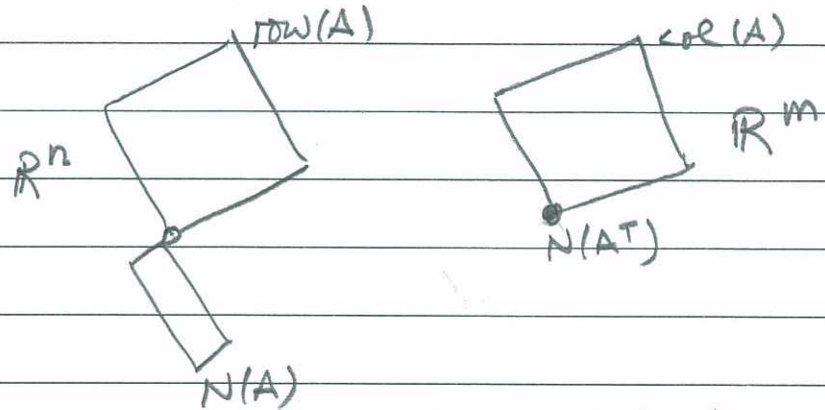
$v_1 \cdots v_r \Rightarrow$ basis row(A)

$u_1 \cdots u_r \Rightarrow$ basis Col(A)

$u_{r+1} \cdots u_m \Rightarrow$ basis $N(A^T)$

We have only r singular values & rest of diag. is zero's.

Let's study the case when $A_{m\times n}$ and $\text{rank}(A) = r = m < n$



$$A\,[v_1 \cdots v_r \; v_{r+1} \cdots v_n] = [u_1 \cdots u_r] \begin{bmatrix} \sigma_1 & & 0 & 0 \\ 0 & & \sigma_r & 0 \end{bmatrix}$$

$$\underset{m\times n}{A} = \underset{m\times m}{U} \; \underset{m\times n}{\Sigma} \; \underset{n\times n}{V^T}$$

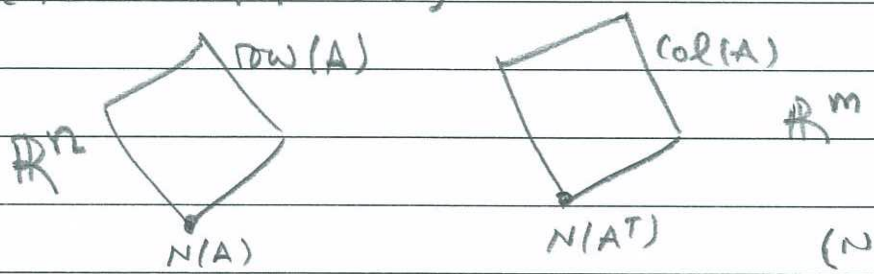
$v_1 \cdots v_r \Rightarrow$ basis row(A)

$u_1 \cdots u_r \Rightarrow$ basis Col(A)

$v_{r+1} \cdots v_n \Rightarrow$ basis N(A)

$\sigma_1 \cdots \sigma_r$ are singular values.

Let's study the case when A has $\text{rank} = r = m = n$ (A is invertible).



(No Null spaces)

$$A\,[v_1 \cdots v_r] = [u_1 \cdots u_r] \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}$$

We can see that the SVD diagonalization works for any matrix A. SVD factorizes A to two $\perp$ matrices (i.e. U & V) and a diagonal matrix (i.e. Σ).

if $A = U \Sigma V^T$

$$A^TA = (U\Sigma V^T)^T U \Sigma V^T$$

Symmetric

$$= V \Sigma^T \underbrace{U^T U}_{I} \Sigma V^T \quad \text{since } U \text{ is } \perp$$

$$= V \Sigma^T \Sigma V^T$$

$= V^{-1}$

which means that V contains the eigenvectors of A^TA, and $(\Sigma^T\Sigma)$ contains the eigenvalues of A^TA.

diagonal

$$\begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_r^2 \end{bmatrix}$$

To find U:

$$AA^T = U\Sigma V^T (U \Sigma V^T)^T$$

$$= U \Sigma \underbrace{V^T V}_{I} \Sigma^T U^T$$

Symmetric

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

which means that U contains the eigenvectors of AA^T, and $(\Sigma\Sigma^T)$ contains the eigenvalues of AA^T

diagonal

$$\begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_r^2 \end{bmatrix}$$

Example: $A = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}$ rank=2 (invertible)

$A^TA = \begin{bmatrix} 25 & 7 \\ 7 & 25 \end{bmatrix}$ $\qquad$ $AA^T = \begin{bmatrix} 32 & 0 \\ 0 & 18 \end{bmatrix}$

we use it to find V $\qquad$ we use it to find U

Find the eigenvalues of A^TA (which are the same as eigenvalues of $AA^T \Rightarrow \lambda_1 = 18$, $\lambda_2 = 32$)

Now, find eigenvectors of A^TA:

$$V = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \quad \text{(note it's to be } \perp \text{)}$$

$$A^TA = V(\Sigma^T\Sigma)V^T$$
$$= \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 18 & 0 \\ 0 & 32 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Now find $U \rightarrow$ Find eigenvectors of AA^T

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

$$\therefore AA^T = U(\Sigma\Sigma^T)U^T$$
$$= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 18 & 0 \\ 0 & 32 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Now $A = U\Sigma V^T$
$$= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \sqrt{18} & 0 \\ 0 & \sqrt{32} \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

A originally had complex eigenvalues and complex eigenvectors. Using SVD we were able to write A as a product of two orthonormal matrices and a diagonal matrix.

Example. $A = \begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix}$ find SVD of A.

rank (A) = 1 (singular)

$$A\,[\underbrace{v_1}_{\text{will come from basis of row(A)}} \;\; \underbrace{v_2}_{\text{basis of N(A)}}] = [\underbrace{u_1}_{\text{basis of col(A)}} \;\; \underbrace{u_2}_{\text{basis of } N(A^T)}]\,\Sigma$$

$\text{eig}(A^TA) = 0, 125 \Rightarrow (\Sigma^T\Sigma) = \begin{bmatrix} 125 & 0 \\ 0 & 0 \end{bmatrix}$

$U = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{bmatrix}$, $\Sigma = \begin{bmatrix} \sqrt{125} & 0 \\ 0 & 0 \end{bmatrix}$, $V^T = \begin{bmatrix} 0.8 & 0.6 \\ 0.6 & -0.8 \end{bmatrix}$