7. Symmetric matrices and SVD

7.1. Symmetric matrices

Def. A matrix $A \in \mathbb{R}^{n\times n}$ is symmetric if $\overbrace{A = A^T}^{\iff a_{ij} = a_{ji} \ \forall i,j}$.

A matrix $A \in \mathbb{R}^{n\times n}$ is anti-symmetric if $\underbrace{A = -A^T}_{\iff a_{ij} = -a_{ji} \ \forall i,j}$

Ex. $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, $\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$, $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

symmetric, symmetric, neither, anti-symm., symmetric and anti-symm.

Ex. All diagonal matrices are symmetric.

All anti-symm. matrices have only zeros on the diagonal

Rmk (i) The symmetric $n\times n$ matrices are a subspace of $\mathbb{R}^{n\times n}$ and its dimension is $\frac{(n+1)n}{2}$ (HW 13)

(ii) If A is symmetric and invertible then A^{-1} is symmetric:

Proof: A invertible $\Rightarrow$ A^{-1} exists

$$(A^{-1})^T \underset{\substack{\uparrow \\ \text{computation rules for} \\ \text{inverses and transposes}}}{=} (A^T)^{-1} \underset{\substack{\uparrow \\ A \text{ symmetric} \\ \text{so } A^T = A}}{=} A^{-1} \Rightarrow A^{-1} \text{ is symmetric}$$

Ex. Diagonalization vs. orthogonal diagonalization

Consider the matrix:

$$A = \begin{pmatrix} -4 & 3 & 0 \\ 3 & 4 & 0 \\ 0 & 0 & 5 \end{pmatrix} \qquad M = \begin{pmatrix} -5 & 1 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

We can find spectrum and eigenspaces as usual

$\sigma_A = \{-5, 5\}$ and

$\left\{ \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} \right\}$ is a basis of $E_A(-5)$

$\left\{ \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \\ 0 \end{pmatrix} \right\}$ is a basis of $E_A(5)$

$\sigma_M = \{-5, 5\}$ and

$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$ is a basis of $E_M(-5)$

$\left\{ \begin{pmatrix} 1 \\ 10 \\ 2 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 5 \\ 0 \end{pmatrix} \right\}$ is a basis of $E_M(5)$

$\Rightarrow$ The matrix has an eigenbasis.

Q: Does it have an orthonormal eigenbasis?

We can make the basis of each eigenspace orthogonal (e.g. by Gram-Schmidt) and then normalize it.

$$\left\{ \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} \right\} \xrightarrow{\text{normalize}} \left\{ \begin{pmatrix} -3/\sqrt{10} \\ 1/\sqrt{10} \\ 0 \end{pmatrix} \right\}$$

$$\left\{ \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \\ 0 \end{pmatrix} \right\} \xrightarrow{\text{orthog}} \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \\ 0 \end{pmatrix} \right\}$$

analogously

normalize $\longrightarrow \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \\ 0 \end{pmatrix} \right\}$

This gives an ONB for each eigenspace:

$\left\{ \begin{pmatrix} -3/\sqrt{10} \\ 1/\sqrt{10} \\ 0 \end{pmatrix} \right\}$ is an ONB of $E_A(-5)$

$\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \\ 0 \end{pmatrix} \right\}$ is an ONB of $E_A(5)$

$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$ is an ONB of $E_M(-5)$

$\left\{ \begin{pmatrix} 1 \\ 10 \\ 2 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 5 \\ 0 \end{pmatrix} \right\}$ is an ONB of $E_M(5)$

Does this yield an orthonormal eigenbasis of the matrix?

Yes! Because each vector in (the basis of) $E_A(-5)$ is orthogonal to each vector in (the basis of) $E_A(5)$.

So the union of the bases of $E_A(-5)$ and $E_A(5)$ forms an ONB of $\mathbb{R}^3$:

$\left\{ \begin{pmatrix} -3/\sqrt{10} \\ 1/\sqrt{10} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \\ 0 \end{pmatrix} \right\}$

No because $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 10 \\ 2 \end{pmatrix} = 1 \neq 0$

So $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 10 \\ 2 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 5 \\ 0 \end{pmatrix} \right\}$ is not an ONB of M.

This cannot be fixed, in particular not by Gram-Schmidt since linear comb. of vectors in $E_M(5)$ and $E_M(-5)$ will not be eigenvectors of M.

e.g. $\begin{pmatrix} 1 \\ 10 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 10 \\ 2 \end{pmatrix}$

Recall : we call a matrix $A \in \mathbb{R}^{n \times n}$ diagonalizable if we can write it as $A = PDP^{-1}$ where $P \in \mathbb{R}^{n \times n}$ invertible and $D \in \mathbb{R}^{n \times n}$ diagonal.

Now observe : If P is also orthogonal , then $P^{-1} = P^T$ and hence $A = PDP^T$.

<u>Def.</u> We say that a matrix $A \in \mathbb{R}^{n \times n}$ is orthogonally diagonalizable if there exists an orthogonal matrix $Q \in \mathbb{R}^{n \times n}$ such that $A = QDQ^T$ and $D \in \mathbb{R}^{n \times n}$ diagonal.

In the above example we saw

- a diagonalization for A

$$A = P \begin{pmatrix} -5 & & \\ & 5 & \\ & & 5 \end{pmatrix} P^{-1} \text{ with } P = \begin{pmatrix} -3 & 1 & 2 \\ 1 & 3 & 6 \\ 0 & 1 & 0 \end{pmatrix}$$

same for M with $P = \begin{pmatrix} 1 & 1 & 1/2 \\ 0 & 10 & 5 \\ 0 & 2 & 0 \end{pmatrix}$

- an orthogonal diagonalization for A

$$A = Q \begin{pmatrix} -5 & & \\ & 5 & \\ & & 5 \end{pmatrix} Q^T \text{ with } Q = \begin{pmatrix} -3/\sqrt{10} & 0 & 1/\sqrt{10} \\ 1/\sqrt{10} & 0 & 3/\sqrt{10} \\ 0 & 1 & 0 \end{pmatrix}$$

does not exist for M

Q: What does symmetry have to do with orthog. diagonalizability?

→ Notice that in the example A was symmetric but M was not.

→ Let us investigate this further...

Assume $A \in \mathbb{R}^{n \times n}$ is a matrix that can be written as $A = QDQ^T$ with Q orthogonal and D diagonal

$$\Rightarrow A^T = (QDQ^T)^T = (Q^T)^T D^T Q^T = QDQ^T = A$$

$\Rightarrow A$ is symmetric.

Lemma 7.0 Let $A \in \mathbb{R}^{n \times n}$. If A is orthogonally diagonalizable

Then A is symmetric.

Thm. 7.1 (The spectral theorem)

Let $A \in \mathbb{R}^{n \times n}$ symmetric. Then:

(1) A has n real eigenvalues (counted with algebraic multiplicity)

(2) $\text{multgeom}_A(\lambda) = \text{multalg}_A(\lambda)$ for all eigenvalues λ

(3) Let v and w be eigenvectors for two distinct eigenvalues of A. Then v and w are orthogonal.

(4) A has an orthonormal eigenbasis.

Lemma 7.0 and Thm. 7.1.(4) give:

Cor 7.2 A matrix $A \in \mathbb{R}^{n \times n}$ is orthogonally diagonalizable if and only if A is symmetric.

Proof of Thm. 7.1

(1) From the fundamental thm of algebra, we know, that A has n eigenvalues (real and complex) counted with algebraic multiplicity. Hence it suffices to show: all eigenvalues of A are real.

Proof: We know from Analysis if $\lambda \in \mathbb{C}$ is a root of a polynomial with real coeff., then also $\bar{\lambda}$ is a root.

Assume $\lambda \in \mathbb{C}$ is eigenvalue of A $\Rightarrow$ λ is a root of $\chi_A(\lambda)$

$\Rightarrow$ $\bar{\lambda}$ is also a root of $\chi_A(\lambda)$ $\Rightarrow$ $\bar{\lambda} \in \mathbb{C}$ is eigenvalue of A

Let v be an eigenvector of A for λ $\Rightarrow$ $Av = \lambda v$

$$\Rightarrow \bar{v}^T A v = \bar{v}^T(\lambda v) = \lambda \bar{v}^T v = \lambda(\bar{v} \cdot v) = \lambda \|v\|^2$$

But also: $A\bar{v} = \overline{Av} = \overline{\lambda v} = \bar{\lambda}\bar{v}$

So $\bar{v}$ is an eigenvector of A for $\bar{\lambda}$

$$\bar{v}^T A v = \bar{v} \cdot Av \overset{\text{HW 13}}{=} A\bar{v} \cdot v = \bar{\lambda}\bar{v} \cdot v = \bar{\lambda}\|v\|^2$$

$$\Rightarrow \lambda \|v\|^2 = \bar{\lambda}\|v\|^2 \text{ and } \|v\|^2 \neq 0 \Rightarrow \lambda = \bar{\lambda} \Rightarrow \lambda \in \mathbb{R}.$$

(2) Omitted

(3) Let $\lambda \neq \mu$ be eigenvalues and v, w respective eigenvectors.

$\Rightarrow Av = \lambda v$, $Aw = \mu w$ and $v \neq 0$, $w \neq 0$

To show: $v \cdot w = 0$

We compute:

$$(\lambda - \mu)(v \cdot w) = \lambda(v \cdot w) - \mu(v \cdot w)$$
$$= (\lambda v) \cdot w - v \cdot (\mu w)$$
$$= Av \cdot w - v \cdot Aw$$
$$= 0 \quad \text{(see HW 13)}$$

$$\Rightarrow \underbrace{(\lambda - \mu)}_{\neq 0}(v \cdot w) = 0 \quad \Rightarrow v \cdot w = 0$$

(4) Let $\mu_1 \dots \mu_k$ be the distinct list of real eigenvalues of A.

Let $\tilde{B}_{\mu_i}$ be a basis of the eigenspace $E_A(\mu_i)$.

For each i, apply Gram-Schmidt to $\tilde{B}_{\mu_i}$ and normalize.

The resulting set B_{μ_i} is an ONB of $E_A(\mu_i)$.

Set $B := \bigcup_{i=1}^{k} B_{\mu_i}$

By part (3): $v \cdot w = 0$ for $v \in B_{\mu_i}$ and $w \in B_{\mu_j}$ when $i \neq j$.

$\Rightarrow B$ is an orthonormal set consisting of eigenvectors of A.

Moreover $\quad \# B = \sum_{i=1}^{k} \# B_{\mu_i} = \sum_{i=1}^{k} \text{geom.mult}_A(\mu_i)$

$$= \underset{\substack{\uparrow \\ \text{Part (2)}}}{} \sum_{i=1}^{k} \text{alg.mult}_A(\mu_i) \underset{\substack{\uparrow \\ \text{Part (1)}}}{=} n$$

$\Rightarrow B$ is an ONB of $\mathbb{R}^n$. $\square$

Geometric interpretation :

Recall that diagonalization corresponds to a change of basis and hence to viewing the linear map $f(x) = Ax$ in terms of a coordinate system. $\rightarrow$ see beginning of Week 11.

In the case of orthogonal diagonalization there is no adjustment of angles in the last step as the n lines given by eigenvectors are already orthogonal.

7.2 The singular value decomposition (SVD)

Q: Can we factorize general matrices $A \in \mathbb{R}^{m\times n}$ in a similar way than symmetric matrices?

Def. Let $A \in \mathbb{R}^{m\times n}$ be a matrix. We will call a product

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

the singular value decomposition of A

if: $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ orthogonal; $\Sigma \in \mathbb{R}^{m\times n}$ "diagonal"

For some $k \le m,n$: $\Sigma_{11}, \dots, \Sigma_{kk} \neq 0$ and all other $\Sigma_{ij} = 0$.

What does this look like?

non-zero entries

$$m\left\{\underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}}_{n}\right. = m\left\{\underbrace{\begin{pmatrix} u_1 & \cdots & u_m \end{pmatrix}}_{m}\right. \underbrace{\overbrace{\begin{pmatrix} * & & & & 0 \\ & * & & & \\ & & \ddots & & \\ & & & * & \\ 0 & & & & 0 \ \ddots \end{pmatrix}}^{k}}_{=n}\Bigg\}m \underbrace{\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}}_{n}\Bigg\}n$$

We will prove that a SVD always exists, how to find U, V and Σ, and in particular what the $*$ are.

Rmk. Orthogonal diagonalization is a (particularly nice) case of a SVD.

In order for the next definition to formally make sense, we need to recall and understand a few things:

Let $A \in \mathbb{R}^{m \times n}$ and consider $A^T A$

$\Rightarrow A^T A \in \mathbb{R}^{n \times n}$

Moreover: $A^T A$ is symmetric: $(A^T A)^T = A^T (A^T)^T = A^T A$

By the spectral thm: $A^T A$ has n real eigenvalues.

Let us list them in descending order $\lambda_1 \geq \ldots \geq \lambda_n$

<u>Claim</u>: all λ_i are ≥ 0

<u>Proof</u>: Let $A^T A v = \lambda v$.

$$\|Av\|^2 = Av \cdot Av = (Av)^T Av = v^T \underbrace{A^T A v}_{= \lambda v}$$

$$= v^T (\lambda v) = \lambda v^T v = \lambda v \cdot v = \lambda \|v\|^2$$

$$\Rightarrow \underbrace{\|Av\|^2}_{\geq 0} = \lambda \underbrace{\|v\|^2}_{\geq 0} \Rightarrow \lambda \geq 0. \qquad \square$$

Def. Let $A \in \mathbb{R}^{m\times n}$. The singular values of A are the

$\sigma_i := \sqrt{\lambda_i}$ where $\lambda_1, \dots, \lambda_n$ are the eigenvalues

of A^TA s.t. $\lambda_1 \geq \lambda_2 \dots \geq \lambda_n \geq 0$ (with multiplicity)

Lemma 7.3 Let $A \in \mathbb{R}^{m\times n}$ $\left(\Rightarrow A^TA \in \mathbb{R}^{n\times n} \text{ symmetric.} \right)$

Let $\{v_1, \dots, v_n\}$ be an ONB of $\mathbb{R}^n$ consisting of eigenvectors

of $A^TA \in \mathbb{R}^{n\times n}$ with corresponding eigenvalues $\lambda_1 \dots \lambda_n$.

We assume we named them ordered by size: $\lambda_1 \geq \dots \geq \lambda_n$.

Assume A has exactly r non-zero singular values.

$\Rightarrow$ $\{Av_1, \dots, Av_r\}$ is an orthogonal basis for $\mathrm{Col}(A)$

and hence $\mathrm{Rank}(A) = r$

Proof We have to show: $\{Av_1, \dots, Av_r\}$ is orthogonal,

does not contain the zero vector, and spans $\mathrm{Col}(A)$.

- $Av_i \cdot Av_j = (Av_i)^T Av_j = v_i^T \underbrace{A^TA v_j}_{=\lambda_j v_j} = v_i^T(\lambda_j v_j) = \lambda_j v_i^T v_j$

  $= \lambda_j v_i \cdot v_j = 0$ if $i \neq j$

  since $v_1 \dots v_n$ is ONB

  $\Rightarrow \{Av_1, \dots, Av_r\}$ is orthogonal

- $$\|Av_j\|^2 = Av_j \cdot Av_j = v_j^T A^T A v_j \overset{\text{as above}}{=} \lambda_j \underbrace{v_j \cdot v_j}_{=1 \text{ since } v_1 \dots v_n \text{ is ONB}}$$

  So $Av_j = 0 \iff \|Av_j\|^2 = 0 \iff \lambda_j = 0$

  By assumption $\lambda_j = 0 \iff j > r$. So $Av_j \neq 0$ for $j = 1 \dots r$.

- Let $y = Ax \in \mathrm{Col}(A)$.

  Let $\alpha_1 \dots \alpha_n \in \mathbb{R}$ such that $x = \alpha_1 v_1 + \dots + \alpha_n v_n$

  $$\Rightarrow Ax = A(\alpha_1 v_1 + \dots + \alpha_n v_n) = \alpha_1 Av_1 + \dots + \alpha_n Av_n$$

  $$= \underbrace{\alpha_1 Av_1 + \dots + \alpha_r Av_r}_{\in \mathrm{Span}\{Av_1, \dots, Av_n\}} + \alpha_{r+1} \underbrace{Av_{r+1}}_{=0} + \dots + \alpha_n \underbrace{Av_n}_{=0}$$

  since $\|Av_i\|^2 = \lambda_i \overset{i>r}{=} 0$

  $$\Rightarrow y = Ax = \alpha_1 Av_1 + \dots + \alpha_r Av_r \in \mathrm{Span}\{Av_1, \dots, Av_r\} \quad \square$$

Lemma 7.3 schematically:

r = # non-zero eigenvalues of A^TA

$v_1 \dots v_r$ are the eigenvectors of these non-zero eigenvalues $\lambda_1 \dots \lambda_r$.

$$\mathbb{R}^n \xrightarrow{A\cdot} \mathbb{R}^m$$

$\underbrace{v_1 \dots v_r}, v_{r+1}, \dots v_n$ ONB $\qquad\qquad$ $Av_1, \dots Av_r$ orthog basis

Recall

Let $A \in \mathbb{R}^{m \times n}$

- An SVD of A is :

$$A = \underset{\text{orthogonal}}{U} \, \Sigma \, \underset{\text{orthogonal}}{V^T} \qquad \Sigma = \begin{pmatrix} * & & \\ & \ddots & \\ & & * \end{pmatrix} \in \mathbb{R}^{m \times n}$$

- $A^T A$ symmetric $\Rightarrow$ has n real eigenvalues λ_i and an ON eigenbasis $v_1 \dots v_n$ where each v_i is an eigenvector for λ_i of $A^T A$!

- The eigenvalues of $A^T A$ are all ≥ 0

- A singular value of A is the square root of a non-zero eigenvalue of $A^T A$: $\sigma_i = \sqrt{\lambda_i} \neq 0$.

- Lemma 7.3 (schematically):

$$r = \#\text{ non-zero eigenvalues of } A^T A : \lambda_1 \geq \dots \geq \lambda_r \geq \overbrace{\lambda_{r+1} \dots \lambda_n}^{\text{all} = 0}$$

$v_1 \dots v_n$ are orthonormal eigenvectors of these eigenvalues $\lambda_1 \dots \lambda_n$.

$$\mathbb{R}^n \xrightarrow{\quad A\cdot \quad} \mathrm{Ran}(A) \subseteq \mathbb{R}^m$$

$v_1 \dots v_r, v_{r+1}, \dots v_n$ ONB $\qquad$ $Av_1, \dots Av_r$ orthog basis of $\mathrm{Ran}(A)$

Thm 7.4 (SVD) Let $A \in \mathbb{R}^{m\times n}$ with $\mathrm{Rank}(A) = r$

Let $\sigma_1 \geq \ldots \geq \sigma_r > 0$ the singular values of A

Then there exist $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ orthogonal such that $A = U \Sigma V^T$

where
$$\Sigma = \begin{pmatrix} \sigma_1 & & & & 0 \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ 0 & & & & \ddots \end{pmatrix} \in \mathbb{R}^{n\times n}$$

(This Thm. says that every matrix has a SVD !!!)

SVD schematically: $A = U \Sigma V^T$

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "A"] \arrow[d, "V^T"'] & \mathbb{R}^m \\ \mathbb{R}^n \arrow[r, "\Sigma"] & \mathbb{R}^m \arrow[u, "U"'] \end{tikzcd}$$

orthogonal change of basis $V^T = V^{-1}$

orthogonal change of basis

matrix that resembles a diagonal matrix

<u>Proof</u> Let $\{v_1, \dots, v_n\}$ be an orthonormal eigenbasis of $A^T A \in \mathbb{R}^{n \times n}$

where each v_i is the eigenvector for λ_i and :

$\lambda_1 \geq \dots \geq \lambda_n \geq 0$. (Recall that all eigenvalues of symmetric matrices are real and non-negative)

Define $V = (v_1, \dots, v_n)$ $\Rightarrow$ V is orthogonal matrix

Let $r :=$ # non-zero eigenvalues λ_i of $A^T A$

Define $\sigma_i := \sqrt{\lambda_i}$ for $i \leq r$ and define : $\Sigma = \begin{pmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ 0 & & & 0 \\ & & & & \ddots \end{pmatrix} \in \mathbb{R}^{m \times n}$

Lemma 7.3 $\Rightarrow$ $\{Av_1, \dots, Av_r\}$ is an orthogonal basis of $Col(A)$
(and rank $A = r$)

Define $u_i := \frac{1}{\|Av_i\|} Av_i$

$\Rightarrow$ $\{u_1, \dots, u_r\}$ is an ONB of $Col(A)$. Recall: $Col(A) \subseteq \mathbb{R}^m$ subspace

Extend $\{u_1, \dots, u_r\}$ to an ONB of $\mathbb{R}^m$: (basis extension thm.)

$\Rightarrow$ $\{u_1, \dots, u_r, u_{r+1}, \dots, u_m\}$ is an ONB of $\mathbb{R}^m$

Define $U := (u_1, \dots, u_m) \in \mathbb{R}^{m \times m}$ $\Rightarrow$ U is an orthogonal matrix

<u>Claim</u> : $AV = U\Sigma$

If we can prove the claim, we are done since:

V orthogonal matrix $\Rightarrow V^T = V^{-1}$

$$\Rightarrow AV = U\Sigma \underset{\cdot V^T}{\Rightarrow} A\underbrace{VV^T}_{=V^{-1}\text{ since } V \text{ is orthogonal square matrix}} = U\Sigma V^T \Rightarrow A = U\Sigma V^T$$

Proof of claim:

$AV = (Av_1, \dots, Av_n)$ by def. of matrix mult.

$$U\Sigma = (u_1 \dots u_m)\begin{pmatrix} \sigma_1 & & & & 0 \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ 0 & & & & \ddots \end{pmatrix} = (\sigma_1 u_1, \dots, \sigma_r u_r, 0, \dots, 0)$$

Recall: $u_i = \frac{Av_i}{\|Av_i\|}$ ($i \le r$) so $Av_i = \|Av_i\| u_i$ *

$$\|Av_i\|^2 = Av_i \cdot Av_i = (Av_i)^T Av_i = v_i^T \overbrace{A^T A v_i}^{=\lambda_i v_i} = v_i^T(\lambda v_i)$$

$$= \lambda_i (v_i^T v_i) = \lambda_i (v_i \cdot v_i) = \underbrace{\lambda_i}_{\substack{=0 \text{ if } i>r \\ =\sigma_i^2 \text{ if } i \le r}} \overbrace{\|v_i\|^2}^{=1 \text{ cause ONB}}$$

$$\Rightarrow \|Av_i\| = \begin{cases} \sigma_i & \text{if } i \le r \\ 0 & \text{if } i > r \end{cases}$$

$$\overset{*}{\Rightarrow} Av_i = \begin{cases} \sigma_i u_i & \text{if } i \le r \\ 0 & \text{if } i > r \end{cases}$$

$\Rightarrow AV = U\Sigma$

$\square$

The proof of Thm. 7.4 provides a recipe for the computation of the SVD of a matrix :

(1) Orthogonal diagonalization of A^TA :

Compute the eigenvalues of the symmetric matrix A^TA and arrange them in descending order : $\lambda_1 \geq \dots \geq \lambda_n$

Find eigenvectors $v_1, \dots, v_n$ of $\lambda_1 \dots \lambda_n$ so that $\{v_1 \dots v_n\}$ ONB

$\rightsquigarrow V := (v_1, \dots, v_n) \in \mathbb{R}^{n \times n}$ orthogonal matrix

(2) Singular values of A :

Let $r :=$ # non-zero eigenvalues λ_i

$\sigma_i := \sqrt{\lambda_i}$ for $i \leq r$

$$\rightsquigarrow \Sigma := \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \\ & & \end{pmatrix} \in \mathbb{R}^{m \times n}$$

(3) Range of A :

Compute $Av_1 \dots Av_r$ (they are orthogonal basis of $\mathrm{Ran}(A)$ by Lem. 7.3)

Normalize : $u_i := \frac{1}{\|Av_i\|} Av_i$ for $i \leq r$

$\rightsquigarrow \{u_1 \dots u_r\}$

(4) Extend basis :

extend $\{u_1 \dots u_r\}$ to a basis of $\mathbb{R}^m$ and make sure that basis is orthonormal.

$\rightsquigarrow \{u_1 \dots u_r, u_{r+1}, \dots, u_m\}$

$\rightsquigarrow U := (u_1, \dots, u_m) \in \mathbb{R}^{m \times m}$ orthogonal

$\underline{\text{Ex.}}$ $A = \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix}$

(1) $A^T A = \begin{pmatrix} 9 & -9 \\ -9 & 9 \end{pmatrix} \Rightarrow \chi_{A^T A}(\lambda) = (\lambda - 18)\lambda$

The two eigenvalues of $A^T A$ are : $0, 18$

List them in descending order : $\lambda_1 = 18, \lambda_2 = 0$

Find eigenvectors of $A^T A$ for λ_1 and λ_2 by the usual methods : $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is an eigenvector for $\lambda_1 = 18$

$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is an eigenvector for $\lambda_2 = 0$

since the eigenvalues are distinct, the eigenvectors are already orthogonal.

Normalize them : $v_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$, $v_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$

Define $V := \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$

(2) $r = 1 \Rightarrow$ singular values are : $\sigma_1 = \sqrt{18}$

$\Sigma := \begin{pmatrix} \sqrt{18} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$

(3) Compute Av_1 : $Av_1 = \begin{pmatrix} 2/\sqrt{2} \\ -4/\sqrt{2} \\ 4/\sqrt{2} \end{pmatrix}$ (recall $r=1$)

Normalize: $u_1 := \frac{1}{\|Av_1\|} Av_1 = \begin{pmatrix} 1/3 \\ -2/3 \\ 2/3 \end{pmatrix}$

(4) Extend $\{u_1\}$ to an ONB of $\mathbb{R}^3$:
We need to find two vectors $u_2, u_3 \in \mathbb{R}^3$
so that $\{u_1, u_2, u_3\}$ ONB of $\mathbb{R}^3$.

Strategy 1 Guess two vectors to $\tilde{u}_2, \tilde{u}_3$ so that $\{u_1, \tilde{u}_2, \tilde{u}_3\}$ indep.

e.g. $\tilde{u}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\tilde{u}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

Then apply Gram-Schmidt to $\{u_1, \tilde{u}_2, \tilde{u}_3\}$ and normalize.

$\rightsquigarrow \{u_1, u_2, u_3\}$.

Strategy 2 Guess two vectors to $\tilde{u}_2, \tilde{u}_3$ orthogonal to u_1

e.g. $\tilde{u}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$, $\tilde{u}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

Apply Gram-Schmidt to $\{\tilde{u}_2, \tilde{u}_3\}$ and normalize!

$\rightsquigarrow \{u_1, u_2, u_3\}$.

more systematic version of strategy 2

Strategy 3 Find vectors $\tilde{u}_2, \tilde{u}_3$ that are orthogonal to u_1 by finding solutions u of the equation $u_1 \cdot u = 0$:

$$\underbrace{\begin{pmatrix} 1/3 \\ -2/3 \\ 2/3 \end{pmatrix}}_{u_1} \cdot \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{u} = 0 \quad \Leftrightarrow \quad \frac{1}{3}x_1 - \frac{2}{3}x_2 + 2/3\, x_3 = 0$$

(linear system!)

Two solutions of that system are e.g. $\tilde{u}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$, $\tilde{u}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

Apply Gram-Schmidt to $\{\tilde{u}_2, \tilde{u}_3\}$ and normalize :

$$u_2 = \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{pmatrix}, \quad u_3 = \begin{pmatrix} -2/\sqrt{45} \\ 4/\sqrt{45} \\ 5/\sqrt{45} \end{pmatrix}$$

Define $U := \begin{pmatrix} 1/3 & 2/\sqrt{5} & -2/\sqrt{45} \\ -2/3 & 1/\sqrt{5} & 4/\sqrt{45} \\ 2/3 & 0 & 5/\sqrt{45} \end{pmatrix}$

(5)
- Check that your result is correct by computing $U \Sigma V^T$ and verify that it is indeed $= A$.
- Check that the columns of U are orthonormal and that the columns of V are orthonormal.