7.1 a) $$\Sigma = \begin{array}{c|c} & \begin{array}{cc} \longleftrightarrow_{r} & \longleftrightarrow_{m-r} \end{array} \\ \hline \begin{array}{c} r \\ \\ \\ n-r \\ \\ \end{array} & \begin{pmatrix} \sigma_1 & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \\ & & \sigma_r & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & & \vdots \\ \vdots & & \vdots & & \ddots & \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} \end{array}$$

Hence :

$$U\Sigma = \begin{array}{c} \begin{array}{cc} \overset{r}{\longleftrightarrow} & \overset{n-r}{\longleftrightarrow} \end{array} \\ \begin{pmatrix} | & & | & & | \\ u_1 & \cdots & u_r & \cdots & u_n \\ | & & | & & | \end{pmatrix} \end{array} \begin{array}{c} r \\ n-r \end{array} \left(\begin{array}{c|c} \tilde{\Sigma} & (0) \\ \hline (0) & (0) \end{array} \right) = \tilde{U} \cdot \left(\tilde{\Sigma} \mid (0) \right)$$

Similarly :

$$\left(\tilde{\Sigma} \mid (0) \right) \cdot V^T = \begin{array}{c} \begin{array}{cc} \overset{r}{\longleftrightarrow} & \overset{m-r}{\longleftrightarrow} \end{array} \\ \left(\begin{array}{c:c} \tilde{\Sigma} & (0) \end{array} \right) \end{array} \begin{array}{c} r \\ m-r \end{array} \begin{pmatrix} \text{---} & v_1 & \text{---} \\ & \vdots & \\ \text{---} & v_r & \text{---} \\ & \vdots & \\ \text{---} & v_m & \text{---} \end{pmatrix}$$

$$= \tilde{\Sigma} \tilde{V}^T.$$

Conclusion: $A = U \Sigma V^T = \tilde{U} (\tilde{\Sigma} \mid 0) V^T$
$= \tilde{U} \tilde{\Sigma} \tilde{V}^T.$

b) We know that $\text{rank}(A) = r$, hence by the rank-nullity theorem: $\dim \text{Ker}(A) = m - r$.

- $v_{r+1}, \dots, v_m$ is orthonormal because $v_1 \dots v_m$ is orthonormal (since the v_i's are the columns of the orthogonal matrix V.

For $i \in \{r+1, \dots, m\}$, $\tilde{V}^T v_i = \begin{pmatrix} \langle v_1, v_i \rangle \\ \vdots \\ \langle v_r, v_i \rangle \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$

because $v_1 \dots v_m$ is orthonormal.

Consequently: $A v_i = \tilde{U} \tilde{\Sigma} \tilde{V}^T v_i = 0$: $v_i \in \operatorname{Ker} A$.

<u>Conclusion</u>: $(v_{r+1}, \dots v_m)$ is an orthonormal family of $m-r$ vectors of $\operatorname{Ker} A$. Since $\dim \operatorname{Ker} A = m-r$, it is an orthonormal basis of $\operatorname{Ker} A$.

- $u_1, \dots u_r$ is orthonormal, because they are columns of the orthogonal matrix U.

Let $(e_1 \dots e_m)$ be the canonical basis of $\mathbb{R}^m$.

Let $i \in \{1, \dots, r\}$. $V^T v_i = e_i$ because $v_1 \dots v_n$ orthonormal and $V = \begin{pmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{pmatrix}$

Let $(e'_1, \dots e'_n)$ be the canonical basis of $\mathbb{R}^n$.

$$\Sigma V^T v_i = \Sigma e_i = \sigma_i e'_i .$$

Hence $A v_i = U \Sigma V^T v_i = \sigma_i U e'_i = \sigma_i u_i$.

Since $i \in \{1, \dots r\}$, we have $\sigma_i \neq 0$, hence $u_i = \frac{1}{\sigma_i} A v_i \in Im(A)$.

- Conclusion: $(u_1, \dots u_r)$ is an orthonormal family of r vectors of $Im(A)$. Since $\dim Im(A) = r$, it is an orthonormal basis of ~~KerA~~ $Im(A)$.

7.2 a) By the spectral theorem (that we can apply to M because M is symmetric) there exists an orthonormal basis $(v_1, \dots, v_n)$ of $\mathbb{R}^n$ consisting of eigenvectors of M: For $i = 1, \dots n$, $M v_i = \lambda_i v_i$, for some $\lambda_i \in \mathbb{R}$.

We have to show that $\lambda_i > 0$ for all i.

Let $i \in \{1, \dots n\}$. Compute $v_i^T M v_i = v_i^T (\lambda_i v_i)$

$$= \lambda_i v_i^T v_i$$

$$= \lambda_i$$

because $v_i^T v_i = \|v_i\|^2 = 1$ since $(v_1 \dots v_n)$ is orthonormal.

By assumption, $v_i^T M v_i > 0$ (because $v_i \neq 0$).

Hence $\lambda_i = v_i^T M v_i > 0$.

Let $P = \begin{pmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{pmatrix}$ and $D = \mathrm{Diag}(\lambda_1 \dots \lambda_n)$. We know that $M = PDP^T$.

P is orthogonal because $(v_1 \dots v_n)$ is orthonormal: P and P^T are thus invertible.

From what we proved in Homework 3: $\mathrm{rank}(M) = \mathrm{rank}(PD)$
$= \mathrm{rank}(D) = n$

because D is diagonal and all the diagonal elements of D are non-zero.

M is $n \times n$ and has rank n: it is invertible.

b) M is symmetric, let $\lambda_1 \geqslant \lambda_2 \geqslant \dots \geqslant \lambda_n$ be its n ordered eigenvalues.

We know that $\lambda_n = \min\limits_{\substack{x \in \mathbb{R}^n \\ \|x\|=1}} x^T M x$.

(Indeed, $-\lambda_n$ is the largest eigenvalue of the symmetric matrix $(-M)$ hence $-\lambda_n = \max\limits_{\|x\|=1} x^T(-M)x = -\min\limits_{\|x\|=1} x^T M x$)

Hence for all non-zero $x \in \mathbb{R}^n$, $\left(\frac{x}{\|x\|}\right)^T M \left(\frac{x}{\|x\|}\right) \geqslant \lambda_n$.

This gives $x^T M x \geqslant \lambda_n \|x\|^2$.

Let $\alpha = |\lambda_n| + 1$.

$$x^T (M + \alpha Id_n) x = x^T M x + \alpha x^T x$$
$$\geqslant \lambda_n \|x\|^2 + (|\lambda_n| + 1) \|x\|^2$$

$$\geqslant \|x\|^2 > 0 \text{ for } x \in \mathbb{R}^n \setminus \{0\}.$$

<u>Conclusion</u> : $M + \alpha Id_n$ is positive definite.

7.4. Let $v_1 \dots v_n$ be an orthonormal family of eigenvectors of M, associated respectively with $\lambda_1, \dots \lambda_n$.

Let $P = \begin{pmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{pmatrix}$ and $D = \mathrm{Diag}(\lambda_1 \dots \lambda_n)$.

We know that $M = PDP^T$.

Let $d \leqslant n$.
Let $U \in \mathbb{R}^{n \times d}$ such that $U^T U = Id_d$.
Define $R = P^T U$. and notice that $R^T R = U^T P P^T U$
$$= U^T Id_n U = Id_d$$

$$Tr(U^T M U) = Tr(U^T P D P^T U)$$
$$= Tr(R^T D R)$$
$$= \sum_{i=1}^{d} (R^T D R)_{i,i}$$

$$(R^T D R)_{i,i} = \sum_{j=1}^{n} (R^T)_{i,j} (DR)_{j,i}$$

$$= \sum_{j=1}^{n} R_{j,i} \sum_{k=1}^{n} D_{j,k} R_{k,i}$$

$$= \sum_{j=1}^{n} R_{j,i} \lambda_j R_{j,i} \quad \text{because } D_{j,k} = \begin{cases} \lambda_j & \text{if } k=j \\ 0 & \text{otherwise.} \end{cases}$$

$$= \sum_{j=1}^{n} \lambda_j R_{j,i}^2 .$$

Hence $Tr(U^T M U) = \sum_{i=1}^{d} \sum_{j=1}^{n} \lambda_j R_{j,i}^2$

$$= \sum_{j=1}^{n} \lambda_j . \underbrace{\left(\sum_{i=1}^{d} R_{j,i}^2 \right)}_{\overset{\text{def}}{=} w_j}$$

- We have $\sum_{j=1}^{n} w_j = d$. Indeed:

$$\sum_{j=1}^{n} w_j = \sum_{j=1}^{n} \sum_{i=1}^{d} R_{j,i}^2 = Tr(R^T R) = Tr(Id_d) = d.$$

- For all j, $0 \leq w_j \leq 1$. Indeed:

orth. projection matrix

$$0 \leq w_j = e_j^T R R^T e_j = \langle e_j, RR^T e_j \rangle \leq \|e_j\| \, \|RR^T e_j\| \leq 1.$$

Hence we have n weights $w_1 \dots w_n$ that sum to d and that are all between 0 and 1.

Hence $\sum_{j=1}^{n} \lambda_j w_j \leq \sum_{j=1}^{d} \lambda_j$ because $\lambda_1 \geq \dots \geq \lambda_n$.

We get that $Tr(U^T M U) \leq \sum_{j=1}^{d} \lambda_j$.

Since U was arbitrarily chosen, we get that

$$\max_{\substack{U \in \mathbb{R}^{n \times d} \\ U^T U = Id_d}} Tr(U^T M U) \leq \sum_{j=1}^{d} \lambda_j .$$

Let us now prove the converse inequality.
Let $U = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_d \\ | & & | \end{pmatrix} \in \mathbb{R}^{n \times d}$.

$v_1 \dots v_d$ are orthonormal, hence $U^T U = Id_d$.

Compute $P^T U = \begin{pmatrix} — v_1 — \\ \vdots \\ — v_n — \end{pmatrix} \begin{pmatrix} | & & | \\ v_1 & \cdots & v_d \\ | & & | \end{pmatrix} = \begin{pmatrix} 1 & \ddots & (0) \\ (0) & \ddots & 1 \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$

Hence $Tr(U^T M U) = Tr\left((P^T U)^T D \, P^T U \right)$

$$= Tr\left(\left(\underbrace{\begin{matrix} 1 & \ddots & (0) \\ (0) & \ddots & 1 \end{matrix}}_{d} \middle| \; (0) \right) \cdot \begin{pmatrix} \lambda_1 & \ddots & (0) \\ (0) & \ddots & \lambda_n \end{pmatrix} \cdot \begin{pmatrix} 1 & \ddots & (0) \\ (0) & \ddots & 1 \\ \hline & (0) & \end{pmatrix} \begin{matrix} \updownarrow d \\ \\ \end{matrix} \right)$$

$= \sum_{j=1}^{d} \lambda_j$. which proves the converse equality.