1.1. Since $A \in \mathbb{R}^{m \times n}$ as a real $m \times n$ matrix.

To prove that the eigenvalues of $AA^T$ and $A^TA$ are real and non-negative, we need to prove that $AA^T$ and $A^TA$ are positive semidefinite

- The product of any matrix and its transpose is symmetric, so both $AA^T$ and $A^TA$ are symmetric matrix. By properties of symmetric matrix that it has real eigenvalues, we know that
- Let $\lambda$ be an eigenvalue of $AA^T$ and $q$ be the eigenvector associated with $\lambda$. Then,

$$(AA^T)q = \lambda q$$
$$q^T(AA^T)q = \lambda q^T q$$
$$\lambda = \frac{q^T(AA^T)q}{q^Tq} = \frac{z^Tz}{q^Tq} \quad \text{where } z = A^Tq$$

Since $z^Tz \geq 0$ and $q^Tq \geq 0$, $\lambda \geq 0$, which implies that all eigenvalues of $AA^T$ are nonnegative.

- Similar process can be done for $A^TA$.

Let $\lambda'$ be an eigenvalue of $A^TA$ and $q'$ be the eigenvector associated with $\lambda'$. Then,

$$(A^TA)q' = \lambda' q'$$
$$q'^T(A^TA)q' = \lambda' q'^T q'$$
$$\lambda' = \frac{(Aq')^T Aq'}{q'^Tq'} = \frac{z'^Tz'}{q'^Tq'} \quad z' = Aq'$$

Since $z'^Tz' \geq 0$ and $q'^Tq' \geq 0$, $\lambda' \geq 0$, which implies that all eigenvalues of $A^TA$ are nonnegative.

Therefore, overall, the eigenvalues of $AA'$ and $A^TA$ are real and non-negative.

1.2. Claim: If $A \in \mathbb{R}^{m\times n}$ as a real $m\times n$ matrix, then $A^TA$ and $AA^T$ have the same nonzero eigenvalues.

Proof: Suppose $A$ is an $m\times n$ matrix, and suppose that $\lambda$ is a nonzero eigenvalue of $A^TA$. Then there exists a nonzero vector $x \in \mathbb{R}^n$ such that $(A^TA)x = \lambda x$ - - - - - - - - - (1)

Multiplying both sides of the equation by $A$:

$$A(A^TA)x = A\lambda x$$
$$(AA^T)(Ax) = \lambda(Ax)$$

Since $\lambda \neq 0$ and $x \neq \vec{0}$, $\lambda x \neq \vec{0}$, and thus, by equation (1), $(A^TA)x \neq \vec{0}$; thus $A^T(Ax) \neq \vec{0}$, which implies $Ax \neq \vec{0}$.

Therefore, $Ax$ is an eigenvector of $AA^T$ corresponding to eigenvalue $\lambda$.

An analogues argument can be used to show that every nonzero eigenvalue of $AA^T$ is an eigenvalue of $A^TA$, and thus completing the proof.

1.3. $A \in \mathbb{R}^{m\times n}$ as a real $m\times n$ matrix.

$$\underline{A} = [\underset{|}{\overset{|}{a_1}}, \underset{|}{\overset{|}{a_2}}, \dots, \underset{|}{\overset{|}{a_n}}] = \begin{bmatrix} - x_1^T - \\ - x_2^T - \\ \vdots^T \\ - x_m - \end{bmatrix} = \underset{m\times m}{U}\ \underset{m\times n}{\Sigma}\ \underset{n\times n}{V^T}$$

1) let $\underline{B} = AA^T = \sum_{i=1}^{n} \underline{a_i}\ \underline{a_i^T}$ as a symmetric matrix

$$\underline{B} = u\Sigma \underbrace{V^TV}_{=1}\Sigma^Tu^T = u\Sigma\Sigma^Tu^T = \underline{u}\begin{bmatrix} \sigma_1^2 & & \underline{0} & \\ & \ddots & & \\ & & \sigma_n^2 & \\ \underline{0} & & & 0 \ddots 0 \end{bmatrix}\underline{u}^T,$$

the form of eigendecomposition of B

Therefore, the left singular vector of $A \Longleftrightarrow$ eigenvectors of $B = AA^T$,

and $\lambda_i = \begin{cases} \sigma_i^2 & i = 1, 2, \dots, n \\ 0 & i = n+1, \dots, m \end{cases}$, which means the

eigenvalues correspond to the square of singular values.

2) Let $\underline{B'} = A^T A = \sum_{i=1}^{m} \underline{x_i} \, \underline{x_i^T}$ as a symmetric matrix

$$\underline{B'} = V \Sigma^T \underbrace{U^T U}_{= I} \Sigma V^T = V \Sigma^T \Sigma V^T = \underline{V} \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{bmatrix} \underline{V^T}$$

Therefore, right singular value of $A \Leftrightarrow$ eigenvectors of $B' = A^T A$, and $\lambda_i = \sigma_i^2$, $i = 1, 2, \cdots, n$, which means the eigenvalues are given by the square of singular values.

In all, this set of eigenvalues are correspond to the square of singular values.

2.1 Let $A^{n \times n}$, a real square matrix.

If rows of $A$ are orthonormal, $AA^T = I$

To prove that columns of $A$ are also orthonormal, we need to show that $A^T A = I$.

Started by $AA^T = I$, multiply each side by $A^{-1}$

$$\underbrace{A^{-1} A}A^T = A^{-1} I$$

$$A^T = A^{-1} I$$

$$A^T = A^{-1}$$

By the properties of inverse matrix that

$AA^{-1} = I = A^{-1}A$, we can substitute $A^{-1}$ by $A^T$.

Therefore, $AA^T = I = A^T A$.

$A^T A = I$ implies that the columns of the matrix are orthonormal.

2.2. • $Ax=b$ is consistent $\Rightarrow$ rank(A) = rank (A|b)

The system $Ax=b$ is consistent means that $b \in$ column space of A, so the dimension of the column space of A = the dimension of column space of A|b.

Since dim( column space of A) = rank (A|b),

rank (A) = rank (A|b)

• rank (A) = rank(A|b) $\Rightarrow$ $Ax=b$ is consistent

rank (A) = rank (A|b)

$\Rightarrow$ dim( column space of A) = dim (column space of A|b).

$\Rightarrow$ a basis in column space of A is a basis in column space of A|b.

Therefore, b is in the span of the columns of A, which means that $Ax=b$ is consistent.

• Geometrically Rank (A) = Rank(A|b) means that adding the vector b as a column to A does not increase the dimension of the vector space spanned by the columns of A.

2.3. $M = [0,1,2]^{1\times3}$

a) $MM^T = [0,1,2]\begin{bmatrix}0\\1\\2\end{bmatrix} = 0+1+4 = 5. = \underset{U}{[1]}\,\underset{\Sigma^2}{[5]}\,\underset{U^T}{[1]}$

Singular value $\Sigma = \sqrt{5}$

$M^TM = \begin{bmatrix}0\\1\\2\end{bmatrix}[0,1,2] = \begin{bmatrix}0&0&0\\0&1&2\\0&2&4\end{bmatrix}$

$$\det(M^TM-\lambda I) = \det\begin{bmatrix} -\lambda & 0 & 0 \\ 0 & 1-\lambda & 2 \\ 0 & 2 & 4-\lambda \end{bmatrix} = \lambda^3 - 5\lambda^2$$

$$\lambda_1 = \lambda_2 = 0 \quad \lambda_3 = 5$$

$$M^TM - 5I = \begin{pmatrix} -5 & 0 & 0 \\ 0 & -4 & 2 \\ 0 & 2 & -1 \end{pmatrix} \qquad V_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad V_2 = \begin{pmatrix} 0 \\ -2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix} \quad V_3 = \begin{pmatrix} 0 \\ 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$$

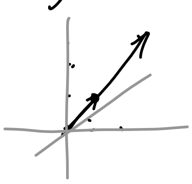
$$M^TM - 0I = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 4 \end{pmatrix} \qquad \Rightarrow V = \begin{pmatrix} 0 & 1/\sqrt{5} & 2/\sqrt{5} \\ 1 & 0 & 0 \\ 0 & -2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix}$$

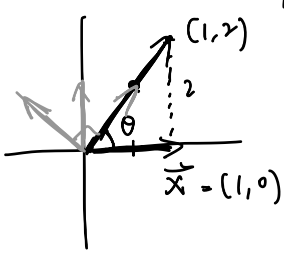
$$\Sigma = (\sqrt{5}, 0, 0)$$

Therefore, by SVD,

$$\overset{1\times 3}{M = [0,1,2]} = \overset{1\times 1}{[1]} \cdot \overset{1\times 3}{[\sqrt{5}, 0, 0]} \overset{3\times 3}{\begin{bmatrix} 0 & 1/\sqrt{5} & 2/\sqrt{5} \\ 1 & 0 & 0 \\ 0 & -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}}$$

b). Since the column space of $M$ is a one dimensional subspace spanned by row vector $[0,1,2]$, we can view $M$ as a transformation of unit vector $(0,1)$ to $(1,2)$. Instead of applying $(1,2)$ to the unit vector, instead, we do rotation, stretching, and then different rotation. In $\mathbb{R}^2$, to transform $(1,0)$ to $(1,2)$ we need to apply






$$\begin{pmatrix} \cos(\tan^{-1} 2) & -\sin(\tan^{-1} 2) \\ \sin(\tan^{-1} 2) & \cos(\tan^{-1} 2) \end{pmatrix}$$

which is $\begin{pmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix}$ by rotation.

Then, to streth it to what we want as $(1,2)$ we need to multiply the vectors by $\sqrt{5}$ as $\|(1,2)\| = \sqrt{5}\|(1,0)\|$.

Therefore, the strething determines the singular value of $\sqrt{5}$.

Since we do not need an extra rotation, the transformation is done, so just set $U$ as 1.

Finally, map the 2D space back to 3D where $M=[0,1,2]$, we just need to expand the column and row by one for rotation, and column by one for stretching.

Therefore, $U=(1)$, $\Sigma=(\sqrt{5},0,0)$, and $V=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{5} & -2/\sqrt{5} \\ 0 & 2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix}$

$V^T=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{5} & 2/\sqrt{5} \\ 0 & -2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix}$

Since the row of matrix can change,

$V^T$ can be written as $\begin{pmatrix} 0 & 1/\sqrt{5} & 2/\sqrt{5} \\ 1 & 0 & 0 \\ 0 & -2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix}$

to match $A$.

Therefore, $A=(1)(\sqrt{5},0,0)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{5} & 2/\sqrt{5} \\ 0 & -2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix}$.

3.1 $A^{m\times n} x^{n\times 1} = b^{m\times 1}$, $m>n$

$$m\underbrace{\begin{bmatrix} & & \\ & & \end{bmatrix}}_{n} \underbrace{\begin{bmatrix} \\ \\ \end{bmatrix}}_{n}{}^{1} = \underbrace{\begin{bmatrix} \\ \\ \end{bmatrix}}_{m}{}^{1}$$

To minimize the error $\|Ax^*-b\|_2$, we can minimize $\|Ax^*-b\|_2^2$.

$$\|Ax^*-b\|_2^2 = \|b-Ax^*\|_2^2 = (b-Ax^*)^T(b-Ax^*)$$
$$= (b^T - x^{*T}A^T)(b-Ax^*)$$
$$= b^Tb - b^TAx^* - x^{*T}A^Tb + x^{*T}A^TAx^*$$

To find the local minimum, we apply derivative to $\|Ax^*-b\|_2^2$

$$\frac{d\|Ax^*-b\|_2^2}{dx^*} = 0 - b^TA - (A^Tb)^T + 2x^{*T}A^TA$$

The local minimum/maximum exist when derivative equals zero

$$\Rightarrow \cancel{2}x^{*T}A^TA = \cancel{2}b^TA$$

Since $A$ is full rank, $A^TA$ is inversible

$$x^{*T} = b^T A (A^T A)^{-1}$$
$$x^* = (A^T A)^{-1} A^T b$$

To check it's local minimum we need to check second derivative is greater than 0 $\frac{d \|Ax^* - b\|_2^2}{dx^{*2}} = 2A^T A$

Since $A^tA$ is non-negative, so $\frac{d\|Ax^*-b\|_2^2}{dx^{*2}} \geq 0$, which means $x^* = (A^T A)^{-1} A^T b$ is a local minimum.

3.2 Least Square via SVD.

$\Rightarrow$ Find $\min_x \|Ax - b\|_2^2$ $\quad\quad AV = \Sigma U$

$$\|Ax-b\|_2^2 = \|(U\Sigma V^T)x - b\|_2^2$$

Since $U$ is orthogonal matrix, $\|U^T\| = 1$. we can add a $U^T$ in the front and since $AV = \Sigma U \Rightarrow U = AV\Sigma^T$, So

$$= \|U^T(AV\Sigma^T\Sigma V^T)x - b\|_2^2$$
$$= \|U^T(AVV^T x - b)\|_2^2$$
$$= \|\Sigma\underbrace{(V^T x)}_{\Downarrow} - U^T b\|_2^2 \quad\quad \text{Let } V^T x = z$$
$$= \|\Sigma z - U^T b\|_2^2 \quad\quad \text{Let Rank}(A) = r.$$
$$= \sum_{i=1}^{r} (\sigma_i z_i - u_i^T b)^2 + \sum_{i=r+1}^{m} (u_i^T b)^2$$

The optimal solution is given by $z_i = \frac{u_i^T b}{\sigma_i}$, for $i = 1, \cdots, r$

and the object becomes $\sum_{i=r+1}^{m} (u_i^T b)^2 = \min_x (\|Ax - b\|_2)$

Thefore, the actual $x^*$:

$$x^* = V z^*$$
$$x^* = \sum_{i=1}^{r} \left(\frac{u_i^T b}{\sigma_i}\right) v_i$$