MA 8463 - Homework 4

Jacob Kutch

September 30, 2022

4.1

Proposition: $||A||_2 = \sigma_1$ where $A \in \mathbb{R}^{mxn}$ and σ_1 is the largest singular value of A.
Proof: Let $A = U\Sigma V^T$ be the singular value decomposition of A, with

$$U = [u_1, u_2, ..., u_n] \quad for \quad U \in \mathbb{R}^{mxn}, \quad s.t. \quad U^T U = I$$

$$\Sigma = diag(\sigma_1, \sigma_2, ..., \sigma_n)$$

$$V = [v_1, v_2, ..., v_n] \quad for \quad V \in \mathbb{R}^{nxn}, \quad s.t. \quad V^T V = I$$

Since the columns of V are the orthonormal basis vectors in $\mathbb{R}^n$ of the orthogonal matrix V,

$$v_i^2 = 1 \quad for \quad i = 1, 2, ..., n \tag{1.1}$$

Additionally, let v_i be eigenvectors of $A^T A$ with eigenvalues $\lambda_i = \sigma_i^2$ and thus by definition of an eigenvalue,

$$A^T A v_i = \lambda_i v_i = \sigma_i^2 v_i \quad for \quad v_i \in \mathbb{R}^n \tag{1.2}$$

Let $x \in \mathbb{R}^n$ be a linear combination of these eigenvectors s.t. $||x||_2 = 1$ where v_i is a basis of $\mathbb{R}^n$. It can be expressed as

$$x = c_1 v_1 + c_2 v_2 + ... + c_n v_n \quad for \quad c_i \in \mathbb{R} \tag{1.3}$$

By (1.2) and (1.3), this means

$$||x||_2^2 = c_1^2 v_1^2 + c_2^2 v_2^2 + ... + c_n^2 v_n^2 = 1 \tag{1.4}$$

Furthermore, considering (1.1), the equation becomes

$$||x||_2^2 = c_1^2 + c_2^2 + ... + c_n^2 = 1 \tag{1.5}$$

Now, considering $||Ax||^2$,

$$||Ax||_2^2 = (Ax)^T (Ax) = xA^T Ax \tag{2.1}$$

From (1.2) and (1.3) again, this becomes

$$xA^T Ax = x(\sigma_i^2 x) = \sigma_i^2 ||x||_2^2 = c_1^2 \sigma_1^2 + c_2^2 \sigma_2^2 + ... + c_n^2 \sigma_n^2 \tag{2.2}$$

Since σ_1 is the largest singular value of A,

$$||Ax||_2^2 = c_1^2 \sigma_1^2 + c_2^2 \sigma_2^2 + ... + c_n^2 \sigma_n^2 \leq \sigma_1^2 (c_1^2 + c_2^2 + ... + c_n^2) \longrightarrow ||Ax||_2^2 \leq \sigma_1^2 ||x||_2^2 \tag{2.3}$$

Finally, when considering (1.5), (2.2), and (2.3),

$$||Ax||_2^2 = ||A||_2^2 ||x||_2^2 \leq \sigma_1^2 ||x||_2^2 \longrightarrow ||A||_2^2 \leq \sigma_1^2 \tag{2.4}$$

If we consider the case where $c_1 = 1$ and $c_i = 0 \quad \forall \quad i \neq 1$, then the inequality in (2.3) becomes equal. For such a choice of c_i (such that (1.5) remains true),

$$||A||_2^2 = \sigma_1^2$$

□

4.2

Proposition: For the Frobenius matrix norm $||A||_F = (\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2)^{1/2}, \quad A \in \mathbb{R}^{mxn}$, we may say for nonzero singular values of A, σ_j, that $||A||_F = (\sigma_1^2 + \sigma_2^2 + ... + \sigma_k^2)^{1/2}$.
Proof: To begin, let $A = U\Sigma V^T$ be the singular value decomposition of A, with

$$U = [u_1, u_2, ..., u_n] \quad for \quad U \in \mathbb{R}^{mxn}, \quad s.t. \quad U^TU = I$$

$$\Sigma = diag(\sigma_1, \sigma_2, ..., \sigma_n)$$

$$V = [v_1, v_2, ..., v_n] \quad for \quad V \in \mathbb{R}^{nxn}, \quad s.t. \quad V^TV = I$$

Note that the Frobenius matrix norm definition may also be written as

$$||A||_F = (Trace(A^TA))^{1/2} \tag{1}$$

where the Trace operation consists of summing up elements along the matrix argument's diagonal, which for A^TA implies a sum of squares of the diagonal elements. This is necessary to note the norm-preserving properties of orthogonal matrices such as U and V. For any orthogonal matrix U, $Trace(U) = Trace(U^T)$ since the diagonal elements will always be the same. A similar reasoning also reveals that for any orthogonal matrix U,

$$||UA||_F^2 = Trace((UA)^T(UA)) = Trace(A^TU^TUA) = Trace(A^TA) = ||A||_F^2. \tag{2}$$

Doing many of the same steps for the full singular value decomposition of A shows

$$||A||_F^2 = ||U\Sigma V^T||_F^2 = Trace((U\Sigma V^T)^T(U\Sigma V^T)) = Trace(V(U\Sigma)^TU\Sigma V^T) = Trace(V\Sigma^TU^TU\Sigma V^T) \tag{3}$$

Since $U^TU = I$,

$$Trace(V\Sigma^TU^TU\Sigma V^T) = Trace(V\Sigma^T\Sigma V^T) = Trace((\Sigma V^T)^T(\Sigma V^T)) = ||\Sigma V^T||_F^2 \tag{4}$$

Since Σ is a diagonal matrix, $\Sigma V^T = V^T\Sigma$. Expanding the Frobenius norm of $V^T\Sigma$ instead gives

$$||V^T\Sigma||_F^2 = Trace((V^T\Sigma)^T(V^T\Sigma)) = Trace(\Sigma^TVV^T\Sigma) = Trace(\Sigma^T\Sigma) = ||\Sigma||_F^2 \tag{5}$$

since $VV^T = I$ for the orthogonal matrix V. Since equations (3)-(5) mean that $||A||_F^2 = ||\Sigma||_F^2$, it also implies that

$$||A||_F = ||\Sigma||_F = (\sum_{j=1}^{n}(\sigma_j)^2)^{1/2} = (\sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2)^{1/2} \tag{6}$$

Since $\sigma_j = 0 \quad \forall j > k = rank(A)$, this may be simplified further as

$$||A||_F = (\sigma_1^2 + \sigma_2^2 + ... + \sigma_k^2)^{1/2} \tag{7}$$

4.3

Proposition: Let $A \in \mathbb{R}^{nxn}$ be symmetric. Then
(a) all the eigenvalues of A are real and
(b) eigenvectors corresponding to distinct eigenvalues are orthogonal

(a)

Proof **:** Assume that $\exists \lambda \in \mathbb{C}$ such that

$$Av = \lambda v \tag{1}$$

where $v \in \mathbb{C}^n$ is an eigenvector corresponding to λ. Since $Av = \lambda v$ consists of vectors of complex numbers that are equal element-wise, their complex conjugates are also equal element-wise. This leads to the equation

$$A\bar{v} = \bar{\lambda}\bar{v} \tag{2}$$

where elements of $\bar{v}$ are the complex conjugates of elements in v and $\bar{\lambda}$ is the complex conjugate of λ. Left multiplying both sides of (1) by $\bar{v}^T$ then yields

$$\bar{v}^T Av = \bar{v}^T(\lambda v) \longrightarrow \bar{v}^T Av = \lambda(\bar{v}^T v) \tag{3}$$

Using a property of matrix transposes, $(AB)^T = B^T A^T$, the left side of (3) may be written as

$$(A^T \bar{v})^T v = \lambda(\bar{v}^T v) \tag{4}$$

Now, noting that $A = A^T$ since A is symmetric and that the resulting form matches that of (2), we may substitute the RHS of (2) and obtain

$$(A\bar{v})^T v = \lambda(\bar{v}^T v) \longrightarrow (\bar{\lambda}\bar{v})^T v = \lambda(\bar{v}^T v) \longrightarrow (\bar{\lambda} - \lambda)(\bar{v}^T v) = 0 \tag{5}$$

Since v (and therefore also $\bar{v}$) is an eigenvector in C and therefore must be nonzero, or equivalently contain a nonzero element, $\bar{v}^T v$ must be nonzero. It must also become a vector in $\mathbb{R}$ since any complex number $a + bi$ multiplied by its conjugate, $a - bi$ becomes a real number $a^2 + b^2$. Therefore, for (5) to hold true,

$$\bar{\lambda} - \lambda = 0 \longrightarrow \bar{\lambda} = \lambda \tag{6}$$

A complex number cannot be equivalent to its conjugate, so λ must be real.

(b)

Proof **:** Let $A \in \mathbb{R}^{nxn}$ be symmetric such that $A = A^T$ and let v_1 and v_2 be eigenvectors of A with corresponding eigenvalues λ_1 and λ_2 such that $\lambda_1 \neq \lambda_2$. Note that for any orthogonal vectors a and b, the inner product is $< a, b >= 0$. Also note that for any inner product of vectors of the form $B \in \mathbb{R}^{nxn}$ and $x, y \in \mathbb{R}^n$,

$$< Bx, y >= y^T Bx \tag{7}$$

For the given eigenvectors and eigenvalues, we have

$$Av_1 = \lambda_1 v_1 \tag{8}$$

$$Av_2 = \lambda_2 v_2 \tag{9}$$

If the eigenvectors $< v_1, v_2 >= 0$, it must also be true that $< Av_1, v_2 >= 0$. Equivalently, $v_2^T Av_1 = 0$
To test this equality, consider that A being symmetric implies

$$v_2^T Av_1 = v_2^T A^T v_1 = (Av_2)^T v_1 \tag{10}$$

From (9), we then know that

$$(Av_2)^T v_1 = (\lambda_2 v_2)^T v_1 = \lambda_2 v_2^T v_1 \tag{11}$$

Also, using (8) in $v_2^T Av_1$ will give

$$v_2^T(Av_1) = v_2^T(\lambda_1 v_1) = \lambda_1 v_2^T v_1 \tag{12}$$

Therefore, combining (11) and (12) produces the equality

$$v_2^T Av_1 = \lambda_1 v_2^T v_1 = \lambda_2 v_2^T v_1 \longrightarrow \lambda_1 v_2^T v_1 - \lambda_2 v_2^T v_1 = 0 \longrightarrow (\lambda_1 - \lambda_2) v_2^T v_1 = 0 \tag{13}$$

Since $\lambda_1 \neq \lambda_2$, this logically means that $v_2^T v_1 = 0$, or $< v_1, v_2 >= 0$. Thus $v_1 \perp v_2$

4.4

Use Matlab to generate a random matrix $A \in R^{8x6}$ *with rank 4. For example,*

```
A = randn(8,4);
A(:,5:6) = A(:,1:2)+A(:,3:4);
[Q,R] = qr(randn(6));
A = A*Q;
```

Figure 1: Code

```
import numpy as np

A = np.random.rand(8,6)
A[:,4:] = A[:,:2] + A[:,2:4]
Q, R = np.linalg.qr(A)    "linalg": Unknown word.

print("Question 7.5.a:")
print(A)
print("\nIt is difficult to tell numerically that the rank is 4. It is easy to see that the rank is at most 5 since \n" \
      "the 5th and 6th columns are the same. It's often difficult to judge linear dependence by inspection otherwise. \n" \
      "However, because of the way A was created, we know that the 5th column is a linear combination of the 1st and 3rd \n" \
      "columns and that the 6th column is a linear combination of the 2nd and 4th columns. If we know this in advance, \n" \
      "we can tell rank(A)=4 by inspection.")

U, Sigma, V = np.linalg.svd(A)    "linalg": Unknown word.
# reformats all entries of sigma to scientific notation with 4 decimal places; there's almost surely a better way for this
Sigma = np.array([np.format_float_scientific(np.float32(Sigma[i]), precision=4) for i in range(len(Sigma))])
print("\nQuestion 7.5.b:")
print(Sigma)
print("The array is ordered from greatest to least singular value. While the first 4 are relatively close together, \n" \
      "the last 2 are much smaller, tending toward 0.")

print("\nQuestion 7.5.c:")
print("rank(A) = ", np.linalg.matrix_rank(A))    "linalg": Unknown word.

print("\nQuestion 7.5.d:")
print("Using TOL = machine precision, \nrank(A) = ", np.linalg.matrix_rank(A, tol = -np.inf))    "linalg": Unknown word.
```

Figure 2: Output

```
Question 7.5.a:
[[0.8029841  0.83172185 0.10844819 0.57429394 0.91143229 1.40601579]
 [0.0358844  0.80861633 0.11015467 0.76526911 0.14603907 1.57388544]
 [0.03568921 0.32465081 0.64791333 0.49878134 0.68360254 0.82343215]
 [0.50785401 0.6963632  0.22349837 0.9625856  0.73135238 1.6589488 ]
 [0.39351254 0.06548446 0.80574161 0.00825468 1.19925415 0.07373914]
 [0.71334131 0.97466807 0.54184883 0.96410575 1.25519014 1.93877382]
 [0.8616648  0.73506915 0.39951364 0.71655264 1.26117844 1.45162179]
 [0.08893386 0.60492047 0.37155024 0.71895236 0.46048409 1.32387283]]

It is difficult to tell numerically that the rank is 4. It is easy to see that the rank is at most 5 since
the 5th and 6th columns are the same. It's often difficult to judge linear dependence by inspection otherwise.
However, because of the way A was created, we know that the 5th column is a linear combination of the 1st and 3rd
columns and that the 6th column is a linear combination of the 2nd and 4th columns. If we know this in advance,
we can tell rank(A)=4 by inspection.

Question 7.5.b:
['5.5436e+00' '1.6184e+00' '8.1803e-01' '2.4405e-01' '1.9761e-16'
 '1.6293e-16']
The array is ordered from greatest to least singular value. While the first 4 are relatively close together,
the last 2 are much smaller, tending toward 0.

Question 7.5.c:
rank(A) =  4

Question 7.5.d:
Using TOL = machine precision,
rank(A) =  6
```

4.5

Let $A \in R^{2x2}$ with singular values $\sigma_1 \geq \sigma_2 > 0$. Show the set $\{Ax \quad | \quad ||x||_2 = 1\}$ (the image of the unit circle) is an ellipse in R^2 whose major and minor semiaxes have lengths σ_1 and σ_2 respectively.

Define the singular values by the eigenvalues λ_1 and λ_2 from the characteristic polynomial $det(A^TA - \lambda I) = 0$. Let v_1 and v_2 be orthonormal eigenvectors of A^TA such that $||v_1|| = 1$ and $||v_2|| = 1$.
If Ax is the image of the unit circle, then x may be the parameterized form of the unit circle, $x = V^Tw$, where $w = \begin{bmatrix} cos(t) \\ sin(t) \end{bmatrix}$ and $V = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$. The proper order of elements in w depends on the direction of the major axis, so for this example, we assume the major axis is the x-axis. If y happened to be the semi-major axis and x the semi-minor axis instead, the only thing that changes within the problem is the order of elements in w. So,

$$x = V^Tw = cos(t)v_1 + sin(t)v_2 \to Ax = cos(t)Av_1 + sin(t)Av_2 \tag{1}$$

We can then expand $||Ax||^2$ as

$$||Ax||^2 = (Ax)^T(Ax) = (cos(t)Av_1 + sin(t)Av_2)^T(cos(t)Av_1 + sin(t)Av_2) = (cos(t)(Av_1)^T + sin(t)(Av_2)^T)(cos(t)Av_1 + sin(t)Av_2)$$

$$= (cos(t)v_1^TA^T + sin(t)v_2^TA^T)(cos(t)Av_1 + sin(t)Av_2)$$

$$= cos^2(t)(v_1^TA^TAv_1) + cos(t)sin(t)(v_1^TA^TAv_2) + sin(t)cos(t)(v_2^TA^TAv_1) + sin^2(t)(v_2^TA^TAv_2)$$

Using the eigenvalue problems $A^TAv_1 = \lambda_1 v_1$ and $A^TAv_2 = \lambda_2 v_2$, we may rewrite this as

$$cos^2(t)(v_1^T\lambda_1 v_1) + cos(t)sin(t)(v_1^T\lambda_2 v_2) + sin(t)cos(t)(v_2^T\lambda_1 v_1) + sin^2(t)(v_2^T\lambda_2 v_2)$$

Since $v_1 \perp v_2$, $\quad v_1^Tv_2 = v_2^Tv_1 = 0$, and because v_1 and v_2 are both unit vectors, it simplifies further as

$$\lambda_1 cos^2(t)(v_1^Tv_1) + \lambda_2 sin^2(t)(v_2^Tv_2) = \lambda_1 cos^2(t)(||v_1||^2) + \lambda_2 sin^2(t)(||v_2||^2) = \lambda_1 cos^2(t) + \lambda_2 sin^2(t)$$

Using the notation $\Sigma = \begin{bmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{bmatrix} = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$, we know that

$$||Ax||^2 = \lambda_1 cos^2(t) + \lambda_2 sin^2(t) = ||\Sigma w||^2 \tag{2}$$

implying that $Ax = \Sigma w = \begin{bmatrix} \sigma_1 cos(t) \\ \sigma_2 sin(t) \end{bmatrix}$, which represents an ellipse in R^2 since the locus of an ellipse is given by $\begin{bmatrix} acos(t) \\ bsin(t) \end{bmatrix}$ where a is the length of the radius of the semi-major axis x, while b is the length of the radius of the semi-minor axis y. In this case, the length of the semi-major axis is the largest singular value of A, σ_1, and the length of the semi-minor axis is the smallest singular value of A, σ_2.

4.6

Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

(a)

Find the SVD of A. You may use the condensed form given in (4.28). To find the singular value decomposition of, we begin by finding AA^T and A^TA. Namely, finding the eigenvalues of AA^T is the first step.

$$AA^T = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 5 & 10 & 15 \\ 10 & 20 & 30 \\ 15 & 30 & 45 \end{bmatrix}, \quad A^TA = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 14 & 28 \\ 28 & 56 \end{bmatrix} \tag{1}$$

Finding the eigenvalues of this with the characteristic equation gives

$$det(AA^T - \lambda I) = 0 \rightarrow det(\begin{bmatrix} 5-\lambda & 10 & 15 \\ 10 & 20-\lambda & 30 \\ 15 & 30 & 45-\lambda \end{bmatrix}) = 0 \rightarrow -\lambda^3 + 70\lambda^2 = 0$$

Thus the eigenvalues of this system, in descending order, are $\lambda_1 = 70$, $\lambda_2 = 0$, and $\lambda_3 = 0$. From this, we know the singular values (and by extension Σ) since $\sigma_i = \sqrt{\lambda_i}$. We may then use these eigenvalues with $A^T A$ to find the columns of V, v_i, using the fact that $v_i \in Null(A^T A - \lambda I)$.

$$rref(\left[\begin{matrix} 14-\lambda_1 & 28 \\ 28 & 56-\lambda_1 \end{matrix} \;\middle|\; \begin{matrix} 0 \\ 0 \end{matrix}\right]) = rref(\left[\begin{matrix} -56 & 28 \\ 28 & -14 \end{matrix} \;\middle|\; \begin{matrix} 0 \\ 0 \end{matrix}\right]) = \left[\begin{matrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{matrix} \;\middle|\; \begin{matrix} 0 \\ 0 \end{matrix}\right]$$

After normalization, this gives eigenvector $v_1 = \begin{bmatrix} \frac{\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} \end{bmatrix}$. Likewise,

$$rref(\left[\begin{matrix} 14-\lambda_2 & 28 \\ 28 & 56-\lambda_2 \end{matrix} \;\middle|\; \begin{matrix} 0 \\ 0 \end{matrix}\right]) = rref(\left[\begin{matrix} 14 & 28 \\ 28 & 56 \end{matrix} \;\middle|\; \begin{matrix} 0 \\ 0 \end{matrix}\right]) = \left[\begin{matrix} 1 & 2 \\ 0 & 0 \end{matrix} \;\middle|\; \begin{matrix} 0 \\ 0 \end{matrix}\right]$$

After normalization, this gives eigenvector $v_2 = \begin{bmatrix} \frac{-2\sqrt{5}}{5} \\ \frac{\sqrt{5}}{5} \end{bmatrix}$. So far, the SVD decomposition has the form

$$A = U\Sigma V^T \rightarrow A = U \begin{bmatrix} \sqrt{70} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{-2\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \end{bmatrix}^T \tag{2}$$

Since V is an orthogonal matrix such that $V^T V = I$,

$$A = U\Sigma V^T \rightarrow AV = U\Sigma \rightarrow Av_i = \sigma_i u_i \rightarrow u_i = \frac{1}{\sigma_i} A v_i \tag{3}$$

$$u_1 = \frac{1}{\sqrt{70}} \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{14}} \\ \frac{2}{\sqrt{14}} \\ \frac{3}{\sqrt{14}} \end{bmatrix}$$

Since A is rank deficient, i.e., $r = rank(A) = 1 < n$, the number of columns of A,

$$A = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 & V_2 \end{bmatrix} = U_1 \Sigma_1 V_1^T \tag{4}$$

Since $\Sigma_1 \in R^{rxr}$ and U_1, V_1 have r columns, the second eigenvector u_2 is unneeded. So we can instead say

$$A = U_1 \Sigma_1 V_1^T = \sqrt{70} \begin{bmatrix} \frac{1}{\sqrt{14}} \\ \frac{2}{\sqrt{14}} \\ \frac{3}{\sqrt{14}} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{-2\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \end{bmatrix}^T$$

(b)

Calculate the pseudoinverse of A, A^+.

$$A^+ = V\Sigma^{-1}U^T = \begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{-2\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{70}} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{14}} & 0 \\ \frac{2}{\sqrt{14}} & 0 \\ \frac{3}{\sqrt{14}} & 0 \end{bmatrix}^T = \begin{bmatrix} \frac{1}{70} & \frac{1}{35} & \frac{3}{70} \\ \frac{1}{35} & \frac{2}{35} & \frac{3}{35} \end{bmatrix}$$

For a rank deficient matrix, the pseudo inverse can be simplified to

$$A^+ = V_1\Sigma_1^{-1}U_1^T = (\frac{1}{\sqrt{70}}) \begin{bmatrix} \frac{\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{14}} \\ \frac{2}{\sqrt{14}} \\ \frac{3}{\sqrt{14}} \end{bmatrix}^T = \begin{bmatrix} \frac{1}{70} & \frac{1}{35} & \frac{3}{70} \\ \frac{1}{35} & \frac{2}{35} & \frac{3}{35} \end{bmatrix} \tag{5}$$

(c)

Calculate the minimum-norm solution of the least-squares problem for the overdetermined system $Ax = b$. Using (5), the minimum-norm solution of the LS problem is

$$x = V_1\Sigma_1^{-1}U_1^T b = \begin{bmatrix} \frac{1}{70} & \frac{1}{35} & \frac{3}{70} \\ \frac{1}{35} & \frac{2}{35} & \frac{3}{35} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{3}{35} \\ \frac{6}{35} \end{bmatrix} \tag{6}$$

(d)

Find a basis for $Null(A)$.
To find $Null(A)$, we perform Gauss-Jordan Elimination on the concatenated matrix

$$[A|0] = \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 3 & 6 & 0 \end{array}\right] = \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

So we find the vector $\begin{bmatrix} -2 \\ 1 \end{bmatrix} \in Null(A)$ making up the basis.

(e)

Find all solutions of the least-squares problem.
All solutions of the LS problem $||Ax - b||_2$ can be written as $x = V_1\Sigma_1^{-1}U_1^T b + V_2 z$ for arbitrary vector z. This equates to the minimum norm solution given in (6) plus another vector $V_2 z$. Since $r = rank(A) = 1$ and V_{r+1} is the first column in the orthonormal basis of $Null(A)$. Thus any linear combination of vectors in the basis of $Null(A)$ may be added to the vector in (6) to give a solution to the LS problem. Formally, all solutions x^* follow

$$x^* = \begin{bmatrix} \frac{3}{35} \\ \frac{6}{35} \end{bmatrix} + z_k \begin{bmatrix} -2 \\ 1 \end{bmatrix} \quad for \quad z_k \in R \tag{7}$$