Homework 8 answers

① Compute the singular-value decompositions of the following matrices. Do not use a computer!

Don't use eigenvalues either.

$$A = \begin{pmatrix} 0 & 0 & 1 & 2 & 3 & 5 \\ 0 & 0 & 3 & 2 & 1 & -2 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$B = \begin{pmatrix} 3 & 0 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & \sqrt{8} & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} 0 & 0 & 1 & 2 & 3 & 5 \\ 0 & 0 & 3 & 2 & 1 & -2 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

We can consider the two blocks separately. The columns in the blue block are orthogonal, whence we get

left s.v.	sing. value	right s.v.
$\frac{1}{2}\begin{pmatrix}1\\1\\1\\1\end{pmatrix}$	2	$\vec{e}_1$
$\frac{1}{2}\begin{pmatrix}1\\1\\-1\\-1\end{pmatrix}$	2	$\vec{e}_2$

The rows in the green block are orthogonal, giving

$\vec{e}_1$	$\sqrt{39}$	$\frac{1}{\sqrt{39}}\begin{pmatrix}1\\2\\3\\5\end{pmatrix}$
$\vec{e}_2$	$3\sqrt{2}$	$\frac{1}{3\sqrt{2}}\begin{pmatrix}3\\2\\1\\-2\end{pmatrix}$
$\vec{e}_3$	$\sqrt{6}$	$\frac{1}{\sqrt{6}}\begin{pmatrix}1\\-2\\1\\0\end{pmatrix}$

Putting these together, and sorting the singular values from largest to smallest, gives

left singular vectors	singular values	right singular vectors
$\vec{e}_1$	$\sigma_1 = \sqrt{39}$	$\frac{1}{\sqrt{39}}\begin{pmatrix}0\\0\\1\\2\\3\\5\end{pmatrix}$
$\vec{e}_2$	$\sigma_2 = 3\sqrt{2}$	$\frac{1}{3\sqrt{2}}\begin{pmatrix}0\\0\\3\\2\\1\\-2\end{pmatrix}$
$\vec{e}_3$	$\sigma_3 = \sqrt{6}$	$\frac{1}{\sqrt{6}}\begin{pmatrix}0\\0\\1\\-2\\1\\0\end{pmatrix}$

$$\sigma_3 = \sqrt{6} \qquad \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix}$$

$$\frac{1}{2}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \qquad \sigma_4 = 2 \qquad \vec{e}_1$$

$$\frac{1}{2}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} \qquad \sigma_5 = 2 \qquad \vec{e}_2$$

$$B = \begin{pmatrix} 3 & 0 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & \sqrt{8} & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

The matrix divides into two blocks.
For each block we have computed the SVD before.

```
MatrixForm /@ SingularValueDecomposition[
```

$$\begin{pmatrix} 3 & 0 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & \sqrt{8} & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}]$$

left ... sing. values ... right

$$\left\{ \begin{pmatrix} -\sqrt{\frac{2}{3}} & -\frac{1}{\sqrt{3}} & 0 & 0 & 0 \\ 0 & 0 & \sqrt{\frac{2}{3}} & -\frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{3}} & -\sqrt{\frac{2}{3}} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} 2\sqrt{3} & 0 & 0 & 0 & 0 \\ 0 & \sqrt{6} & 0 & 0 & 0 \\ 0 & 0 & \sqrt{3} & 0 & 0 \\ 0 & 0 & 0 & \sqrt{3} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \right\}$$

Pseudoinverses

(2) By hand, compute the pseudoinverses of the matrices

$$A = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \end{pmatrix}_{1\times n} \qquad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}_{n\times 1}$$

$$C = \begin{pmatrix} 1 & 2 & 3 & 0 \\ -4 & 2 & 0 & 2 \\ 1 & 1 & -1 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 4 \end{pmatrix}$$

Answer:
The SVDs are

$$A = \sqrt{n} \cdot (1) \cdot \frac{1}{\sqrt{n}}(1 \; 1 \cdots 1) \qquad B = \|\vec{b}\| \cdot \frac{\vec{b}}{\|b\|} \cdot (1)$$

singular values
left sing. vectors
right sing. vectors

$\Rightarrow$ the pseudoinverses are

$$A^+ = \frac{1}{\sqrt{n}} \frac{1}{\sqrt{n}} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} (1) \qquad B^+ = \frac{1}{\|b\|} \cdot (1) \frac{\vec{b}^T}{\|b\|}$$

→ the pseudoinverses are

$$A^+ = \frac{1}{\sqrt{n}}\frac{1}{\sqrt{n}}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}(1) = \boxed{\frac{1}{n}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}_{n\times 1}} \qquad B^+ = \frac{1}{\|b\|}\cdot(1)\,\frac{\vec{b}^T}{\|b\|} = \boxed{\frac{(b_1\ b_2\ \cdots\ b_n)}{b_1^2+\cdots+b_n^2}}$$

$$C = \vec{e}_1\vec{v}_1^T + \vec{e}_2\vec{v}_2^T + \vec{e}_3\vec{v}_3^T$$

where $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is pairwise orthogonal

Its SVD is therefore

$$C = \|v_1\|\,\vec{e}_1\frac{\vec{v}_1^T}{\|v_1\|} + \|v_2\|\,\vec{e}_2\frac{\vec{v}_2^T}{\|v_2\|} + \|v_3\|\,\vec{e}_3\frac{\vec{v}_3^T}{\|v_3\|}$$

$$\Rightarrow C^+ = \frac{1}{\underbrace{\|v_1\|^2}_{=14}}\vec{v}_1\vec{e}_1^T + \frac{1}{\underbrace{\|v_2\|^2}_{=24}}\vec{v}_2\vec{e}_2^T + \frac{1}{\underbrace{\|v_3\|^2}_{=4}}\vec{v}_3\vec{e}_3^T = \begin{pmatrix}\frac{1}{14} & \frac{-4}{24} & \frac{1}{4}\\ \frac{2}{14} & \frac{2}{24} & \frac{1}{4}\\ \frac{3}{14} & 0 & -\frac{1}{4}\\ 0 & \frac{2}{24} & \frac{1}{4}\end{pmatrix} = \begin{pmatrix}\frac{1}{14} & -\frac{1}{6} & \frac{1}{4}\\ \frac{1}{7} & \frac{1}{12} & \frac{1}{4}\\ \frac{3}{14} & 0 & -\frac{1}{4}\\ 0 & \frac{1}{12} & \frac{1}{4}\end{pmatrix}$$

$D = 2e_1e_2^T + \begin{pmatrix}0\\3\\4\end{pmatrix}e_3^T$ is the SVD (up to normalization)

$$\Rightarrow D^+ = \frac{1}{2}\vec{e}_2\vec{e}_1^T + \frac{\vec{e}_3(0\ 3\ 4)}{\|(0\ 3\ 4)\|^2} = \begin{pmatrix}0 & 0 & 0\\ 1/2 & 0 & 0\\ 0 & 3/25 & 4/25\end{pmatrix}$$

③ Is $(AB)^+ = B^+A^+$ always true for pseudoinverses?

No! Give a counterexample.

Answer:

$$A = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} = A^+ \qquad B = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0\\ 1 & 0\end{pmatrix} \Rightarrow B^+ = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}$$

$$\Rightarrow B^+A^+ = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}$$

$$AB = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} \Rightarrow (AB)^+ = \sqrt{2}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} \neq B^+A^+ \checkmark$$

SVD and condition number

④

ⓐ Without computing the norm exactly, argue why

$$100 \le \left\|\begin{pmatrix}1 & -100\\ 0 & 1\end{pmatrix}\right\| \le 101.$$

ⓑ What is the inverse of $A = \begin{pmatrix}1 & -100\\ 0 & 1\end{pmatrix}$?

What is the condition number of A?

What is the condition number of A^{-1}?

ⓒ Find vectors $\vec{b} \in \mathbb{R}^2$ and $\vec{\delta} \in \mathbb{R}^2$ such that $\|\delta\|$ is "small" compared to $\|b\|$, and yet

$$\| A^{-1}(b+\delta) - A^{-1} b \|$$

is "large" compared to $\|A^{-1} b\|$. (Use your own judgment for what should count as small or large.)

(Hint: Compute the SVD and experiment a bit, using Matlab.)

ⓐ We showed in class that $\|A\| \geq \max_{i,j} |a_{ij}|$.
This gives the lower bound of 100.
For the upper bound, use the triangle inequality:

$$\left\| \begin{pmatrix} 1 & -100 \\ 0 & 1 \end{pmatrix} \right\| \leq \left\| \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\| + \left\| \begin{pmatrix} 0 & -100 \\ 0 & 0 \end{pmatrix} \right\| = 1 + 100 \quad \checkmark$$

ⓑ $A^{-1} = \begin{pmatrix} 1 & 100 \\ 0 & 1 \end{pmatrix} \Rightarrow 100 \leq \|A^{-1}\| \leq 101$, also

$\Rightarrow$ condition number $\kappa(A) = \kappa(A^{-1}) = \|A\| \cdot \|A^{-1}\| \in [100^2, 101^2]$

To compute the condition number exactly, note

$$\|A\| = \|A^{-1}\| = \sqrt{5001 + 100\sqrt{2501}}$$

ⓒ $\|A^{-1}(b+\delta) - A^{-1} b\| = \|A^{-1}\delta\|$

A is a rank-2 matrix, so its SVD is

$$A = \lambda_1 u_1 v_1^T + \lambda_2 u_2 v_2^T,$$

for $\{u_1, u_2\}$ and $\{v_1, v_2\}$ certain orthonormal bases for $\mathbb{R}^2$, with $\lambda_1 > \lambda_2 > 0$. (In fact $\lambda_1 \approx 100$, $\lambda_2 \approx 1/100$.)

$$A^{-1} = \frac{1}{\lambda_1} v_1 u_1^T + \frac{1}{\lambda_2} v_2 u_2^T$$

If we want $A^{-1}\delta$ to be large, we should choose δ in the u_2 direction.
If we want $A^{-1}b$ to be small, we should choose b in the u_1 direction.

So try setting

$$\vec{b} = \vec{u}_1, \quad \vec{\delta} = \varepsilon \vec{u}_2 \quad \Rightarrow \frac{\|\delta\|}{\|b\|} = \varepsilon \ \underline{\text{small}}!$$

$$\Rightarrow A^{-1} b = \frac{1}{\lambda_1} \vec{v}_1 \qquad \Rightarrow \|A^{-1} b\| = \frac{1}{\lambda_1} \approx \frac{1}{100}$$

$$A^{-1}\delta = \frac{1}{\varepsilon \lambda_2} \vec{v}_2 \qquad \Rightarrow \|A^{-1}\delta\| = \frac{\varepsilon}{\lambda_2} \approx 100\varepsilon$$

$$\Rightarrow \frac{\|A^{-1}\delta\|}{\|A^{-1}b\|} = \frac{\varepsilon/\lambda_2}{1/\lambda_1} = \frac{\lambda_1}{\lambda_2} \cdot \varepsilon$$

$$= \left(\begin{array}{c}\text{condition}\\ \text{number}\end{array} \frac{\lambda_1}{\lambda_2}\right) \cdot \frac{\|\delta\|}{\|b\|}$$

22
10,000

⑤ ⓐ Let $H_3 = \begin{pmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{pmatrix}$.

Compute $b = H_3 x$ for $x = (1,1,1)$ and $x = (0, 6, -3.6)$.

A small change Δb produces a large change Δx.

ⓑ Compute numerically the largest and smallest singular values of the 7×7 "Hilbert matrix" H_7, a matrix whose (i,j) entry is $\frac{1}{i+j-1}$.

(Hint: Google the Matlab "hilb" command.)

ⓒ If $H_7 x = b$ with $\|b\| = 1$, how large can $\|x\|$ be? If b has roundoff error less than 10^{-16} in norm, how large an error can this cause in x.

ⓐ

```
>> format rat      % display rational approximations of the output
H3 = hilb(3)
format short e  % go back to decimal output
H3 * [1 1 1]'
H3 * [0 6 -3.6]'

H3 =

       1              1/2            1/3
       1/2            1/3            1/4
       1/3            1/4            1/5


ans =

   1.8333e+00
   1.0833e+00
   7.8333e-01


ans =

   1.8000e+00
   1.1000e+00
   7.8000e-01
```

ⓑ

```
>> H7 = hilb(7);
svd(H7)

ans =

   1.6609e+00
   2.7192e-01
   2.1290e-02
   1.0086e-03
   2.9386e-05
   4.8568e-07
   3.4939e-09
```

$\lambda_1 =$ 1.6609e+00 ← largest singular value

$\lambda_7 =$ 3.4939e-09 ← smallest singular value

ⓒ Let the SVD be $H_7 = \sum_{i=1}^{7} \lambda_i u_i v_i^T$, for the singular values λ_i found above.

If $b = u_i$, then $x = \frac{1}{\lambda_7} v_i$ solves $Ax = b$, with $\|b\| = 1$ and $\|x\| = \frac{1}{\lambda_7} \approx 2.86 \times 10^8$

An error in b in the direction of u_i will similarly cause an error in x that is $\frac{1}{\lambda_7}$ times as large. So if the error in b is 10^{-16}, the error in x can be as large as 2.86×10^{-8}.

⑥ Let A be an $m \times n$ real matrix, with singular-value decomposition

$$A = \sum_i \sigma_i \vec{v}_i \vec{u}_i^T.$$

Assume that the singular values are sorted, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k \geq 0$.

a) Give the SVD for the matrix AA^T.
What is its condition number?
Why is $R(AA^T) = R(A)$?
Why is $N(A^TA) = N(A)$?

b) Give the SVD for the $(m+n)\times(m+n)$ matrix

$$B = \begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}.$$

What is its condition number?
In terms of A^+, what is the pseudoinverse B^+?
(Hint: What is B applied to $(\vec{0}, \vec{u}_i) \in \mathbb{R}^{m+n}$?
How about B applied to $(\vec{v}_i, \vec{0}) \in \mathbb{R}^{m+n}$?)

ⓐ $A = \sum_i \sigma_i \vec{u}_i \vec{v}_i^T$

$$\Rightarrow \quad AA^T = \left(\sum_i \sigma_i u_i v_i^T\right)\left(\sum_j \sigma_j v_j u_j^T\right)$$

$$= \sum_{i,j} \sigma_i \sigma_j u_i u_j^T \cdot (v_i \cdot v_j)$$

$$= \sum_i \sigma_i^2 u_i u_i^T \qquad \text{since } v_i \cdot v_j = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}$$

This is an SVD of AA^T since the $\{u_i\}$ form an orthonormal basis. The left and right singular vectors are the same. Note that the singular values of AA^T are the squares of the singular values of A.

The condition number is

$$K(AA^T) = \frac{\text{largest sing. value}}{\text{smallest sing. value}} = \frac{\sigma_1^2}{\sigma_k^2} = K(A)^2$$

$$R(AA^T) = \text{Span}\left(\{\vec{u}_i \mid \sigma_i > 0\}\right) = R(A).$$

ⓑ

(b)

$$B = \begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}, \quad A = \sum_i \sigma_i u_i v_i^T$$

Right singular vectors	Corresponding Left singular vectors	Corresponding Singular values
$(\vec{0}, \vec{v}_i)$	$(\vec{u}_i, \vec{0})$	σ_i
$(\vec{u}_i, \vec{0})$	$(\vec{0}, \vec{v}_i)$	σ_i

This gives a full SVD for B since the claimed right singular vectors are all orthonormal, as are the claimed left singular vectors, and the total number of each is $m+n$ — so we have a complete basis and aren't missing any.

Here's an example in Matlab:

```
>> m = 2; n = 3;
>> A = randn(m, n);
>> [U,D,V] = svd(A)

U =

  -0.5231   -0.8523
  -0.8523    0.5231


D =

   1.3951         0         0
        0    0.7179         0


V =

  -0.3783    0.7489   -0.5441
  -0.0129    0.5834    0.8121
   0.9256    0.3142   -0.2110
```

```
>> [U,D,V] = svd([zeros(m) A; A' zeros(n)])

U =

  -0.0000    0.5231   -0.8523   -0.0000   -0.0000
        0    0.8523    0.5231         0         0
  -0.3783   -0.0000    0.0000   -0.7489    0.5441
  -0.0129   -0.0000    0.0000   -0.5834   -0.8121
   0.9256   -0.0000    0.0000   -0.3142    0.2110


D =

   1.3951         0         0         0         0
        0    1.3951         0         0         0
        0         0    0.7179         0         0
        0         0         0    0.7179         0
        0         0         0         0    0.0000


V =

  -0.5231         0         0    0.8523         0
  -0.8523         0         0   -0.5231         0
        0    0.3783    0.7489         0    0.5441
        0    0.0129    0.5834         0   -0.8121
        0   -0.9256    0.3142         0    0.2110
```

Condition #: $\kappa(B) = \frac{\sigma_1}{\sigma_k} = \kappa(A)$.

Pseudoinverse:

$$B^+ = \begin{pmatrix} 0 & (A^+)^T \\ A^+ & 0 \end{pmatrix}$$

Least-squares regression

7 a) Find the projection of b onto the column space of A:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -2 & 4 \end{bmatrix}, \qquad b = \begin{bmatrix} 1 \\ 2 \\ 7 \end{bmatrix}.$$

a) Find the projection of b onto the column space of A:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -2 & 4 \end{bmatrix}, \qquad b = \begin{bmatrix} 1 \\ 2 \\ 7 \end{bmatrix}.$$

Split b into $p + q$, with p in the column space and q perpendicular to that space. Which of the four subspaces, $R(A), R(A^T), N(A)$ or $N(A^T)$ contains q?

b) Using your answer from part ⓐ, find the $\vec{x} \in \mathbb{R}^2$ that minimizes $\|A\vec{x} - \vec{b}\|$.

Answer: Since q is perpendicular to $R(A)$, $q \in N(A^T)$.

Here's one way of finding p and q in Matlab:

```
>> format rat;  % tell Matlab to output rational approximations to the outputs
A = [1 1; 1 -1; -2 4];
b = [1; 2; 7];
X = orth(A); % performs Gram-Schmidt orthogonalization to get an orthonormal basis for R(A)
P = X * X';  % since the columns of X are orthonormal, call them x_1,...,x_n,
             % X*X' = sum_j x_j x_j' is the projection onto the range of X,
                                                % which equals the range of A

p = P * b
q = b - p

p =

     23/11
    -14/11
     65/11

q =

    -12/11
     36/11
     12/11
```

Here's another way, using that $A \underbrace{A^+}_{\text{pseudo inverse of } A} = P_{R(A)}$, as shown in class:

```
>> P = A * pinv(A);
p = P * b
q = b - p
```

And, finally, this is also easy to do by hand.

1. First find the x that minimizes $\|Ax - b\|$ by solving the normal equations:

$$A^T A x = A^T b$$

$$\Rightarrow \begin{pmatrix} 6 & -8 \\ -8 & 18 \end{pmatrix} x = \begin{pmatrix} -11 \\ 27 \end{pmatrix}$$

$$\Rightarrow x = \frac{1}{22} \begin{pmatrix} 9 \\ 37 \end{pmatrix}$$

2. Applying A then gives $P_{R(A)} b$:

$$P_{R(A)} b = Ax = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ -2 & 4 \end{pmatrix} \cdot \frac{1}{22} \begin{pmatrix} 9 \\ 37 \end{pmatrix} = \boxed{\frac{1}{11} \begin{pmatrix} 23 \\ -14 \\ 65 \end{pmatrix}} \checkmark$$

Another way would be to use Gram-Schmidt to get an orthonormal basis for $R(A)$, then project onto it as in the first Matlab code transcript above.

b) We want to solve $A\vec{x} = P_{R(A)}\vec{b} = \vec{p} = \frac{1}{11}\begin{pmatrix}23\\-14\\65\end{pmatrix}$

All these ways work:

```
>> A \ p

ans =

     9/22
    37/22
```

```
>> pinv(A) * [1 2 7]'

ans =

     9/22
    37/22
```

```
>> (A' * A) \ (A' * [1 2 7]')

ans =

     9/22
    37/22
```

```
>> A \ [1 2 7]'

ans =

     9/22
    37/22
```

8. For $\mathbf{A} \in \Re^{m\times n}$ and $\mathbf{b} \in \Re^{m}$, prove that $\mathbf{x_2}$ is a least squares solution for $\mathbf{Ax} = \mathbf{b}$ if and only if $\mathbf{x_2}$ is part of a solution to the larger system

$$\begin{pmatrix}\mathbf{I}_{m\times m} & \mathbf{A}\\ \mathbf{A}^T & \mathbf{0}_{n\times n}\end{pmatrix}\begin{pmatrix}\mathbf{x}_1\\ \mathbf{x}_2\end{pmatrix} = \begin{pmatrix}\mathbf{b}\\ \mathbf{0}\end{pmatrix}.$$

Answer:

Assume x_2 is a least-squares solution to $Ax = b$,

i.e. $A^TAx_2 = A^Tb$.

Let $x_1 = b - Ax_2$

$\Rightarrow \begin{pmatrix} I & A \\ A^T & 0\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix} b \\ A^Tx_1 = 0\end{pmatrix}$ ✓

Conversely, if $\begin{pmatrix}x_1\\x_2\end{pmatrix}$ solves $\begin{pmatrix} I & A \\ A^T & 0\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix} b\\0\end{pmatrix}$

then $x_1 = b - Ax_2$ by the first block of equations

and $A^Tx = 0 = A^Tb - A^TAx_2$ by the 2^{nd} block ✓