EECS 16A  Designing Information Devices and Systems I  
Summer 2022  

Homework 7

This homework is OPTIONAL and will not be graded

1. Reading Assignment

For this homework, please review Note 20 (Op-Amp Current Source and Circuit Design), Note 21 (Inner Products and GPS), Note 22 (Trilateration and Correlation) and Note 23 (Least Squares).

2. Inner Product Properties

**Learning Goal:** The objective of this problem is to exercise useful identities for inner products.

Our definition of the inner product in \( \mathbb{R}^n \) is:
\[
\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \mathbf{x}^\top \mathbf{y}, \quad \text{for any } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n
\]

Prove the following identities in \( \mathbb{R}^n \):

(a) \( \langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle \)

**Solution:** This is seen by direct expansion:

Let \( x_i, y_i \in \mathbb{R} \), then
\[
\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \mathbf{x}^\top \mathbf{y}, \quad \text{for any } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n
\]

(b) \( \langle \mathbf{x}, \mathbf{x} \rangle = ||\mathbf{x}||^2 \)
Solution:
\[
\begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_n
\end{bmatrix},
\begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_n
\end{bmatrix} = x_1 \cdot x_1 + x_2 \cdot x_2 + \cdots + x_n \cdot x_n
\]
\[
= x_1^2 + x_2^2 + \cdots + x_n^2
\]
\[
= (\sqrt{x_1^2 + x_2^2 + \cdots + x_n^2})^2
\]
\[
= (\|\mathbf{x}\|)^2
\]

The inner product of a vector with itself is its norm squared.

(c) \( \langle -\mathbf{x}, \mathbf{y} \rangle = -\langle \mathbf{x}, \mathbf{y} \rangle \).

Solution:
\[
\langle -\mathbf{x}, \mathbf{y} \rangle = \begin{bmatrix}
  -x_1 \\
  -x_2 \\
  \vdots \\
  -x_n
\end{bmatrix} \cdot \begin{bmatrix}
  y_1 \\
  y_2 \\
  \vdots \\
  y_n
\end{bmatrix}
\]
\[
= -x_1 \cdot y_1 - x_2 \cdot y_2 - \cdots - x_n \cdot y_n
\]
\[
= -(x_1 \cdot y_1 + x_2 \cdot y_2 + \cdots + x_n \cdot y_n)
\]
\[
= -\langle \mathbf{x}, \mathbf{y} \rangle
\]

Flipping the sign of one of the vectors in the inner product flips the sign of the inner product, but does not change the magnitude.

(d) \( \langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle \)

Solution:
\[
\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \mathbf{x}^\top (\mathbf{y} + \mathbf{z})
\]
\[
= \mathbf{x}^\top \mathbf{y} + \mathbf{x}^\top \mathbf{z}
\]
\[
= \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle
\]

The inner product is distributive.

(e) \( \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + 2\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \)

Solution:
\[
\langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle
\]
\[
= \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle
\]
\[
= \langle \mathbf{x}, \mathbf{x} \rangle + 2 \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle
\]
3. Cauchy-Schwarz Inequality

**Learning Goal:** The objective of this problem is to understand and prove the Cauchy-Schwarz inequality for real-valued vectors.

The Cauchy-Schwarz inequality states that for two vectors \( \vec{v}, \vec{w} \in \mathbb{R}^n \):

\[
| \langle \vec{v}, \vec{w} \rangle | = | \vec{v}^T \vec{w} | \leq \| \vec{v} \| \cdot \| \vec{w} \|
\]

In this problem we will prove the Cauchy-Schwarz inequality for vectors in \( \mathbb{R}^2 \).

Take two vectors:

\[
\vec{v} = r \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} \quad \text{and} \quad \vec{w} = t \begin{bmatrix} \cos \phi \\ \sin \phi \end{bmatrix}, \quad \text{where} \quad r > 0, \ t > 0, \ \theta, \ \text{and} \ \phi \ \text{are scalars. Make sure you understand why any vector in} \ \mathbb{R}^2 \ \text{can be expressed this way and why it is acceptable to restrict} \ r, t > 0.
\]

(a) In terms of some or all of the variables \( r, t, \theta, \) and \( \phi \), what are \( \| \vec{v} \| \) and \( \| \vec{w} \| \)? **Hint:** Recall the trig identity: \( \cos^2 x + \sin^2 x = 1 \)

**Solution:** We use the trig identity \( \cos^2 x + \sin^2 x = 1 \) to show:

\[
\| \vec{v} \| = \sqrt{v_1^2 + v_2^2} = \sqrt{r^2 \cos^2 \theta + r^2 \sin^2 \theta} = r
\]

Similarly, \( \| \vec{w} \| = t \).

(b) In terms of some or all of the variables \( r, t, \theta, \) and \( \phi \), what is \( \langle \vec{v}, \vec{w} \rangle \)? **Hint:** The trig identity \( \cos(a) \cos(b) + \sin(a) \sin(b) = \cos(a - b) \) may be useful.

**Solution:** We use the trig identity \( \cos(x) \cos(y) + \sin(x) \sin(y) = \cos(x - y) \) to show:

\[
\langle \vec{v}, \vec{w} \rangle = (r \cos \theta)(t \cos \phi) + (r \sin \theta)(t \sin \phi) = r \cdot t \cos(\theta - \phi)
\]

(c) Show that the Cauchy-Schwarz inequality holds for any two vectors in \( \mathbb{R}^2 \). **Hint:** consider your results from part (b). Also recall \(-1 \leq \cos x \leq 1\) and use both inequalities.

**Solution:** We use the fact that \( \cos x \leq 1 \) to show:

\[
\langle \vec{v}, \vec{w} \rangle = r \cdot t \cos(\theta - \phi) = \| \vec{v} \| \| \vec{w} \| \cos(\theta - \phi) \leq \| \vec{v} \| \| \vec{w} \|
\]

We use the fact that \( \cos x \geq -1 \) to show:

\[
\langle \vec{v}, \vec{w} \rangle = r \cdot t \cos(\theta - \phi) = \| \vec{v} \| \| \vec{w} \| \cos(\theta - \phi) \geq -\| \vec{v} \| \| \vec{w} \|
\]
Therefore:

\[-\|\vec{v}\|\|\vec{w}\| \leq \langle \vec{v}, \vec{w} \rangle \leq \|\vec{v}\|\|\vec{w}\|,\]

which gives us that

\[|\langle \vec{v}, \vec{w} \rangle| \leq \|\vec{v}\|\|\vec{w}\|.

(d) Note that the inequality states that the inner product of two vectors must be less than or equal to the product of their magnitudes. What conditions must the vectors satisfy for the equality to hold? In other words, when is \(\langle \vec{v}, \vec{w} \rangle = \|\vec{v}\| \cdot \|\vec{w}\|?\)

Solution:

\[\langle \vec{v}, \vec{w} \rangle = \|\vec{v}\|\|\vec{w}\|\cos(\theta - \phi) = \|\vec{v}\|\|\vec{w}\|\]

We see that the equality holds when the angle between the two vectors is zero. Note that when the angle is zero, the vectors would be linearly dependent.

4. Orthonormal Matrices

Definition: A matrix \(U \in \mathbb{R}^{n \times n}\) is called an orthonormal matrix if \(U^{-1} = U^T\) and each column of \(U\) is a unit vector.

Orthonormal matrices represent linear transformations that preserve angles between vectors and the lengths of vectors. Rotations and reflections, useful in computer graphics, are examples of transformations that can be represented by orthonormal matrices.

(a) Let \(U\) be an orthonormal matrix. For two vectors \(\vec{x}, \vec{y} \in \mathbb{R}^n\), show that \(\langle \vec{x}, \vec{y} \rangle = \langle U\vec{x}, U\vec{y} \rangle\), assuming we are working with the Euclidean inner product.

Solution:

\[\langle U\vec{x}, U\vec{y} \rangle = \|U\vec{x}\|\|U\vec{y}\|\cos(\theta - \phi) = \|\vec{x}\|\|\vec{y}\|\]

\[\cos(\theta - \phi) = 1\]

\[\theta - \phi = 0\]

(b) Show that \(\|U\vec{x}\| = \|\vec{x}\|\), where \(\|\cdot\|\) is the Euclidean norm.

Solution:

\[\|U\vec{x}\| = \sqrt{\langle U\vec{x}, U\vec{x} \rangle} = \sqrt{\langle \vec{x}, \vec{x} \rangle} = \|\vec{x}\|\]

The second equality follows from the identity proved in part (a).

(c) How does multiplying \(\vec{x}\) by \(U\) affect the length of the vector? That is, how do the lengths of \(U\vec{x}\) and \(\vec{x}\) compare? Solution:

Recall that the \(L_2\), or Euclidean norm of a vector is the length. As we proved in part b), \(U\) does not affect the norm of \(\vec{x}\). In other words, the length of \(\vec{x}\) is the same before and after applying \(U\)! This allows us to transform \(\vec{x}\) in ways that may make analysis easier while preserving its length! You will have the opportunity to explore this further in EECS 16B.

5. Mechanical Projections

Learning Goal: The objective of this problem is to practice calculating projection of a vector and the corresponding squared error.
(a) Find the projection of \( \vec{b} = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix} \) onto \( \vec{a} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \). What is the squared error between the projection and \( \vec{b} \), i.e. \( \|e\|^2 = \|\text{proj}_a(\vec{b}) - \vec{b}\|^2 \)?

Solution:

\[
\text{proj}_a(\vec{b}) = \frac{\langle \vec{b}, \vec{a} \rangle}{\|\vec{a}\|^2} \vec{a} = \frac{\vec{b}^T \vec{a}}{\|\vec{a}\|^2} \vec{a}
\]

First, compute \( \|\vec{a}\|^2 = \langle \vec{a}, \vec{a} \rangle = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = 2 \).

Second, compute \( \langle \vec{b}, \vec{a} \rangle = \begin{bmatrix} 3 & 2 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 2 \).

Plugging in, \( \text{proj}_a(\vec{b}) = \frac{2}{2} \vec{a} = \vec{a} \).

The squared error between \( \vec{b} \) and its projection onto \( \vec{a} \) is \( \|e\|^2 = \|\vec{a} - \vec{b}\|^2 = 12 \).

(b) Find the projection of \( \vec{b} = \begin{bmatrix} 1 \\ 4 \\ -5 \end{bmatrix} \) onto the column space of \( A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \). What is the squared error between the projection and \( \vec{b} \), i.e. \( \|e\|^2 = \|\text{proj}_{\text{col}(A)}(\vec{b}) - \vec{b}\|^2 \)?

Solution: Let \( A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \) and \( \bar{x} \in \mathbb{R}^2 \) such that the projection of \( \vec{b} \) onto the column space of \( A \) is \( A\bar{x} \).

We will compute \( \bar{x} \) by solving the following least squares problem,

\[
\min_{\bar{x}} \|A\bar{x} - \vec{b}\|^2
\]

The solution yields,

\[
\hat{x} = (A^T A)^{-1} A^T \vec{b}
\]

\[
= \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}^T \begin{bmatrix} 1 \\ 4 \\ -5 \end{bmatrix}
\]

\[
= \begin{bmatrix} 0.5 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -4 \\ 4 \end{bmatrix}
\]

\[
= \begin{bmatrix} -2 \\ 4 \end{bmatrix}
\]

Plugging in, the projection of \( \vec{b} \) onto the column space of \( A \) is \( \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -2 \\ 4 \\ -2 \end{bmatrix} = \begin{bmatrix} -2 \\ 4 \\ -2 \end{bmatrix} \).

The squared error between the projection and \( \vec{b} \) is \( \|\vec{c}\|^2 = \left| \left| \begin{bmatrix} -2 \\ 4 \\ -2 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right| \right|^2 = 18 \).
6. Mechanical Trilateration

**Learning Goal:** The objective of this problem is to practice using trilateration to find the position based on the distance measurements and known beacon locations.

Trilateration is the problem of finding one’s coordinates given distances from known beacon locations. For each of the following trilateration problems, you are given 3 beacon locations \((\vec{s_1}, \vec{s_2}, \vec{s_3})\) and the corresponding distance \((d_1, d_2, d_3)\) from each beacon to your location.

(a) \(\vec{s_1} = \left[\begin{array}{c} 4 \\ 5 \end{array}\right], \ d_1 = 5, \ \vec{s_2} = \left[\begin{array}{c} 1 \\ -1 \end{array}\right], \ d_2 = 2, \ \vec{s_3} = \left[\begin{array}{c} -11 \\ 6 \end{array}\right], \ d_3 = 13\). First, use any graphing calculator or ipython to graph the set of constraints given by \((\vec{s_1}, \vec{s_2}, \vec{s_3})\) and \((d_1, d_2, d_3)\), and take note of the number of solutions, or possible locations that you could be. Then use trilateration to find your location or possible locations. If a solution does not exist, state that it does not.

**Solution:** From graphing these equations we can see there is a single point of intersection, or in other words, a single possible solution to our location.

Now, we show a general approach to the trilateration problem, so that we can immediately write the linear system of equations for all three parts and solve for our solution algebraically. However, if you solved directly using concrete values, give yourself full credit.

\[
\begin{align*}
||\vec{x} - \vec{s_1}||^2 &= d_1^2 \\
||\vec{x} - \vec{s_2}||^2 &= d_2^2 \\
||\vec{x} - \vec{s_3}||^2 &= d_3^2
\end{align*}
\]

Now, let’s show this algebraically with trilateration. We can expand each left hand side out in terms of the definition of the norm:

\[
\begin{align*}
||\vec{x} - \vec{s_1}||^2 &= (\vec{x} - \vec{s_1})^T (\vec{x} - \vec{s_1}) \\
|\vec{x}^T \vec{x} - 2\vec{x}^T \vec{s_1} + \vec{s_1}^T \vec{s_1}| &= d_1^2 \\
|\vec{x}^T \vec{x} - 2\vec{x}^T \vec{s_2} + \vec{s_2}^T \vec{s_2}| &= d_2^2 \\
|\vec{x}^T \vec{x} - 2\vec{x}^T \vec{s_3} + \vec{s_3}^T \vec{s_3}| &= d_3^2
\end{align*}
\]

Finally, take one equation and subtract it from the other two to get a system of linear equations in \(\vec{x}\):

\[
\begin{align*}
2\vec{x}^T \vec{s_3} - 2\vec{x}^T \vec{s_1} &= d_1^2 - d_2^2 + \vec{s_3}^T \vec{s_3} - \vec{s_1}^T \vec{s_1} \\
2\vec{x}^T \vec{s_3} - 2\vec{x}^T \vec{s_2} &= d_2^2 - d_3^2 + \vec{s_3}^T \vec{s_3} - \vec{s_2}^T \vec{s_2}
\end{align*}
\]

We can express as a matrix equation in \(\vec{x}\):

\[
\begin{pmatrix}
2(\vec{s_3} - \vec{s_1})^T \\
2(\vec{s_3} - \vec{s_2})^T
\end{pmatrix}
\vec{x} =
\begin{pmatrix}
d_1^2 - d_2^2 + ||\vec{s_3}||^2 - ||\vec{s_1}||^2 \\
d_2^2 - d_3^2 + ||\vec{s_3}||^2 - ||\vec{s_2}||^2
\end{pmatrix}
\]

We have that:

\[
\begin{align*}
2(\vec{s_3} - \vec{s_1}) &= \begin{pmatrix} -30 \\ 2 \end{pmatrix} \\
2(\vec{s_3} - \vec{s_2}) &= \begin{pmatrix} -24 \\ 14 \end{pmatrix}
\end{align*}
\]

\[
\begin{align*}
d_1^2 - d_2^2 + ||\vec{s_3}||^2 - ||\vec{s_1}||^2 &= 25 - 169 + 157 - 41 = -28 \\
d_2^2 - d_3^2 + ||\vec{s_3}||^2 - ||\vec{s_2}||^2 &= 4 - 169 + 157 - 2 = -10
\end{align*}
\]
Which gives us the system\[
\begin{bmatrix}
-30 & 2 \\
-24 & 14
\end{bmatrix}
\begin{bmatrix}
x \\
y
\end{bmatrix}
= \begin{bmatrix}
-28 \\
-10
\end{bmatrix}
\]
with solution \(x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}\).

A solution existing for this system of linear equations does not necessarily guarantee consistency of the system of nonlinear equations, but we can validate:

\[
\begin{align*}
\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 4 \\ 5 \end{bmatrix} \|^2 &= \| \begin{bmatrix} -3 \\ -4 \end{bmatrix} \|^2 = 25 = d_1^2 \\
\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ -1 \end{bmatrix} \|^2 &= \| \begin{bmatrix} 0 \\ 2 \end{bmatrix} \|^2 = 4 = d_2^2 \\
\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} -11 \\ 6 \end{bmatrix} \|^2 &= \| \begin{bmatrix} 12 \\ -5 \end{bmatrix} \|^2 = 169 = d_3^2
\end{align*}
\]

(b) \(\vec{s}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\), \(d_1 = 5\sqrt{2}\), \(\vec{s}_2 = \begin{bmatrix} 10 \\ 0 \end{bmatrix}\), \(d_2 = 5\sqrt{2}\), \(\vec{s}_3 = \begin{bmatrix} 5 \\ 0 \end{bmatrix}\), \(d_3 = 5\). First, use any graphing calculator or ipython to graph the set of constraints given by \((\vec{s}_1, \vec{s}_2, \vec{s}_3)\) and \((d_1, d_2, d_3)\), and take note of the number of solutions, or possible locations that you could be. Then use trilateration to find your location or possible locations. Why can’t we precisely determine our location, even though we have the same number of measurements as part (a)? Can we use our original constraints to narrow down our set of possible solutions we got from trilateration?

**Solution:** Graphing our constraints gives us two points of intersection.

Now, let’s try to algebraically solve for these points using trilateration. Using the linearization approach from part (a) we get:

\[
2(\vec{s}_3 - \vec{s}_1) = \begin{bmatrix} 10 \\ 0 \end{bmatrix}
\]

\[
2(\vec{s}_3 - \vec{s}_2) = \begin{bmatrix} -10 \\ 0 \end{bmatrix}
\]

\[
d_1^2 - d_2^2 + \|\vec{s}_3\|^2 - \|\vec{s}_1\|^2 = 50 - 25 + 25 - 0 = 50
\]

\[
d_2^2 - d_3^2 + \|\vec{s}_3\|^2 - \|\vec{s}_2\|^2 = 50 - 25 + 25 - 100 = -50
\]

Which gives us the system
\[
\begin{bmatrix}
10 & 0 \\
-10 & 0
\end{bmatrix}
\begin{bmatrix}
x \\
y
\end{bmatrix}
= \begin{bmatrix}
50 \\
-50
\end{bmatrix}
\]

with solution \(x = \begin{bmatrix} 5 \\ \alpha \end{bmatrix}\). We can see that by having collinear beacons, we may not be able to precisely determine our location (short exercise: how does this relate to span and vector spaces?)

However, from the graph we know that not all values of \(\alpha\) are valid, so we can plug our solution back into the third distance equation:

\[
\| \begin{bmatrix} 5 \\ \alpha \end{bmatrix} - \begin{bmatrix} 5 \\ 0 \end{bmatrix} \|^2 = 5^2 \implies \alpha^2 = 25 \implies \alpha = \pm 5
\]

The system of nonlinear equations is consistent with this solution. We do not have enough information to uniquely determine our location, but we know we are at either \(x = \begin{bmatrix} 5 \\ 5 \end{bmatrix}\) or \(x = \begin{bmatrix} 5 \\ -5 \end{bmatrix}\).

(c) \(\vec{s}_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}\), \(d_1 = 5\), \(\vec{s}_2 = \begin{bmatrix} 0 \\ -2 \end{bmatrix}\), \(d_2 = 2\), \(\vec{s}_3 = \begin{bmatrix} -12 \\ 5 \end{bmatrix}\), \(d_3 = 12\). First, use any graphing calculator or ipython to graph the set of constraints given by \((\vec{s}_1, \vec{s}_2, \vec{s}_3)\) and \((d_1, d_2, d_3)\), and take note of the number
of solutions, or possible locations that you could be. Then use trilateration to find your location or possible locations. If a solution does not exist, state that it does not.

**Solution:** Graphing our equations gives us no points of intersection, meaning that there will be no solutions.

Now, let’s show this algebraically with trilateration. Using again what was shown in part (a) we have that:

\[
\begin{align*}
2(\vec{s}_3 - \vec{s}_1) &= \begin{bmatrix} -30 \\ 2 \end{bmatrix} \\
2(\vec{s}_3 - \vec{s}_2) &= \begin{bmatrix} -24 \\ 14 \end{bmatrix} \\
\|\vec{s}_1\|^2 - d_1^2 + \|\vec{s}_3\|^2 - \|\vec{s}_1\|^2 &= 25 - 144 + 169 - 25 = 25 \\
\|\vec{s}_2\|^2 - d_2^2 + \|\vec{s}_3\|^2 - \|\vec{s}_2\|^2 &= 4 - 144 + 169 - 4 = 25
\end{align*}
\]

Which gives us the system

\[
\begin{bmatrix} -30 & 2 \\ -24 & 14 \end{bmatrix} \vec{x} = \begin{bmatrix} 25 \\ 25 \end{bmatrix}
\]

While a solution, \(\vec{x} = \begin{bmatrix} -\frac{75}{186} \\ \frac{25}{25} \end{bmatrix}\), for this system of linear equations exists, it will yield inconsistent distances when substituted back into the nonlinear equations.

\[\|\vec{s}_1 - \vec{x}\|^2 = (3 + \frac{75}{93})^2 + (4 - \frac{75}{186})^2 \neq d_1^2\]

Therefore there is no solution.

**7. Mechanical: Linear Least Squares**

![Graph](image)

(a) Consider the above data points. Find the linear model of the form

\[\vec{b} = \vec{a}x\]
that best fits the data, i.e. find the scalar value of $x$ that minimizes the squared error

$$
\|\vec{e}\|^2 = \left\| \begin{bmatrix} b_1 \\ \vdots \\ b_4 \end{bmatrix} - \begin{bmatrix} a_1 \\ \vdots \\ a_4 \end{bmatrix} x \right\|^2 = \|\vec{b} - \vec{a}x\|^2. \tag{7}
$$

**Note:** by using this linear model, we are implicitly forcing the fit equation to go through the origin, i.e. $0 = x_0$ for all $x$.

**Do not use IPython for this calculation and show your work.** Once you’ve computed $x$, compute the squared error between your model’s prediction and the actual $\vec{b}$ values as shown in Equation 7. Plot the best fit line along with the data points to examine the quality of the fit. (It is okay if your plot of $\vec{b} = \vec{a}x$ is approximate.)

**Reminder:** $\hat{x} = (\vec{a}^T \vec{a})^{-1} \vec{a}^T \vec{b}$

**Solution:**

Define $\vec{a} = [2 \ 4 \ 6 \ 8]^T$ and $\vec{b} = [2 \ 6 \ 7 \ 8]^T$. Applying the linear least squares formula, we get

$$
\hat{x} = (\vec{a}^T \vec{a})^{-1} \vec{a}^T \vec{b}
$$

$$
= \begin{pmatrix}
2 & 2 \\
4 & 4 \\
6 & 6 \\
8 & 8 \\
\end{pmatrix}
\begin{pmatrix}
2 \\
4 \\
6 \\
8 \\
\end{pmatrix}
^{-1}
\begin{pmatrix}
2 \\
4 \\
6 \\
8 \\
\end{pmatrix}
\vec{b}
$$

$$
= (120)^{-1} (134) = 1.1167
$$

The error between the model’s prediction and actual $b$ values is

$$
\vec{e} = \vec{b} - \hat{\vec{b}} = \vec{b} - \hat{x} \vec{a}
$$

$$
= \begin{bmatrix}
2 \\
6 \\
7 \\
8 \\
\end{bmatrix}
- 1.1167
\begin{bmatrix}
2 \\
4 \\
6 \\
8 \\
\end{bmatrix}
= \begin{bmatrix}
-0.234 \\
1.534 \\
0.3 \\
-0.934 \\
\end{bmatrix}
$$

and the sum of squared errors is

$$
\vec{e}^T \vec{e} = 3.367
$$
(b) You will notice from your graph that you can get a better fit by adding a $b$-intercept. That is we can get a better fit for the data by assuming a linear model of the form

$$\tilde{b} = x_1 \tilde{a} + x_2.$$ 

In order to do this, we need to augment our $A$ matrix for the least squares calculation with a column of 1’s (do you see why?), so that it has the form

$$A = \begin{bmatrix} a_1 & 1 \\ \vdots & \vdots \\ a_4 & 1 \end{bmatrix}.$$ 

Find $x_1$ and $x_2$ that minimize the squared error

$$\| \tilde{e} \|^2 = \| \begin{bmatrix} b_1 \\ \vdots \\ b_4 \end{bmatrix} - \begin{bmatrix} a_1 & 1 \\ \vdots & \vdots \\ a_4 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \|^2.$$ 

(8)

**Do not use IPython for this calculation and show your work.**

**Reminder:** $\hat{x} = (A^T A)^{-1} A^T \tilde{b}$

**Reminder:**

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Compute the squared error between your model’s prediction and the actual $\tilde{b}$ values as shown in Equation 8. Plot your new linear model. Is it a better fit for the data?

**Solution:**

Let $\bar{x} = [x_1 \\ x_2]^T$. Using the linear least squares formula with the new augmented $A$ matrix, we calculate the optimal approximation of $\bar{x}$ as

$$\bar{x} = (A^T A)^{-1} A^T \tilde{b} = \begin{bmatrix} 2 & 1 \\ 4 & 1 \\ 6 & 1 \\ 8 & 1 \end{bmatrix}^T \begin{bmatrix} 2 & 1 \\ 4 & 1 \\ 6 & 1 \\ 8 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 4 \\ 6 \\ 8 \end{bmatrix} = \begin{bmatrix} 120 \\ 20 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 4 \\ 6 \\ 8 \end{bmatrix} = \begin{bmatrix} 134 \\ 23 \end{bmatrix}$$

$$\hat{x} = [\hat{x}_1 \ \hat{x}_2] = \begin{bmatrix} 0.95 \\ 1 \end{bmatrix}$$

The linear model’s prediction of $\tilde{b}$ is given by

$$\tilde{b} = A \hat{x} = \begin{bmatrix} 2 & 1 \\ 4 & 1 \\ 6 & 1 \\ 8 & 1 \end{bmatrix} \begin{bmatrix} 0.95 \\ 1 \end{bmatrix} = \begin{bmatrix} 2.9 \\ 4.8 \\ 6.7 \\ 8.6 \end{bmatrix}$$
and the error is given by
\[
\vec{e} = \vec{b} - \hat{\vec{b}} = \begin{bmatrix} -0.9 & 1.2 & 0.3 & -0.6 \end{bmatrix}^T
\]

The summed squared error is
\[
\vec{e}^T \vec{e} = 2.7
\]

We can see both qualitatively from the plots and quantitatively from the sum of the squared errors that the fit is better with the \( b \)-intercept.

8. Proof: Least Squares

Let \( \vec{x} \) be the solution to a linear least squares problem.

\[
\vec{x} = \arg\min_{\vec{x}} \| \vec{b} - A\vec{x} \|^2
\]

Show that the error vector \( \vec{b} - A\vec{x} \) is orthogonal to the columns of \( A \) by direct manipulation (i.e. plug the formula for the linear least squares estimate into the error vector and then check if \( A^T \) times the vector is the zero vector.)

Solution:

We want to show that the error in the linear least squares estimate is orthogonal to the columns of the \( A \), i.e., we want to show that \( A^T (\vec{b} - A\vec{x}) \) is the zero vector. Plugging in the linear least squares formula for \( \vec{x} \), we get

\[
A^T (\vec{b} - A\vec{x}) = A^T (\vec{b} - A(A^TA)^{-1}A^T\vec{b}) = A^T\vec{b} - A^T A(\vec{A}^T\vec{A})^{-1}A^T\vec{b} = A^T\vec{b} - \vec{b}
\]