EECS 16A Designing Information Devices and Systems I
Summer 2023 Homework 7

This homework is due Saturday, August 5, 2023, at 23:59.
Self-grades are due Friday, August 11, 2023, at 23:59.

Submission Format
Your homework submission should consist of **one** file.

- `hw7.pdf`: A single PDF file that contains all of your answers (any handwritten answers should be scanned).

Submit the file to the appropriate assignment on Gradescope.

1. Reading Assignment

For this homework, please read Note 20 (Circuit Design), Note 21 (Inner Products and GPS), Note 22 (Trilateration and Correlation), and read Note 23 (Least Squares). You are always encouraged to read beyond this as well.

(a) In trilateration, the distances between the beacons and the unknown location $\vec{x}$ involve quadratic terms of $\vec{x}$. What trick can we use to get a system of linear equations in $\vec{x}$?

Solution: We can get a system of linear equations by subtracting one non-linear equation from another and eliminating the quadratic terms.

(b) Suppose the signal $x[n]$ is only defined for timesteps $0, 1, \cdots, 5$. For the purpose of computing linear cross-correlation, what value of $x[n]$ do we assume when n is a timestep out of the range: $0 \leq n \leq 5$ (e.g. $n = 6$ or $n = -1$)?

Solution: When n is a timestep out of the range: $0 \leq n \leq 5$, we consider $x[n]$ to be zero.

2. Transistor Equivalent

Consider the amplifier circuit in Fig. 1 which amplifies input V_{in} to output V_{out}. The circuit accomplishes this by using a bipolar junction transistor or BJT. The BJT is a three-terminal circuit element with nodes B, C, and E.

In some situations, the BJT can be modeled with an equivalent linear circuit containing a voltage-dependent current source as shown in Fig. 2.

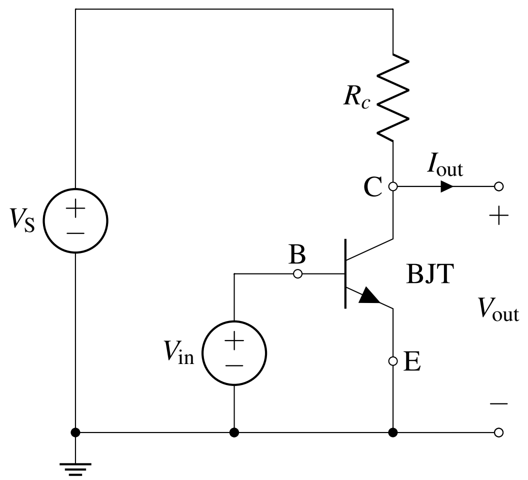


Figure 1: Amplifier circuit with BJT

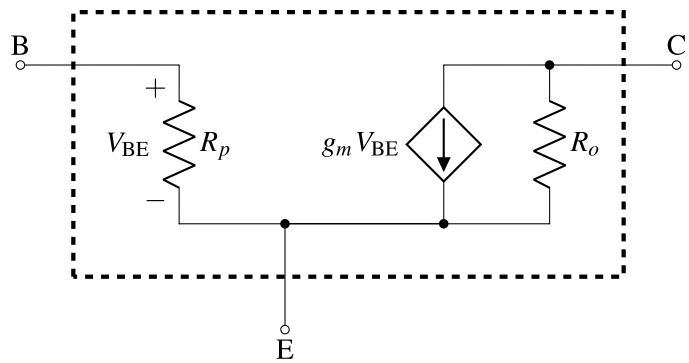


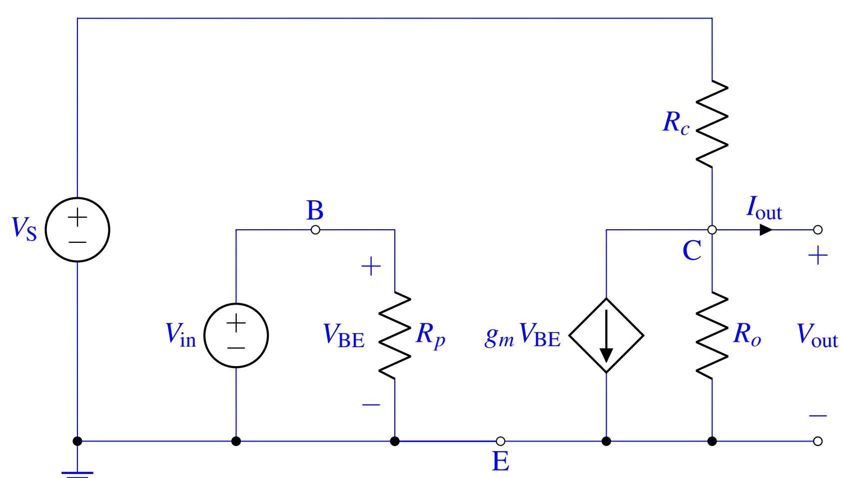
Figure 2: Equivalent circuit model for BJT

We want to find the Thevenin and Norton equivalent of the amplifier circuit across the terminals V_{out} (i.e., between nodes C and E).

Note: You can use the parallel operator (||) in your final answers.

(a) Redraw the original amplifier circuit in Fig. 1 but with the BJT equivalent circuit model in Fig. 2 substituted.

Solution:

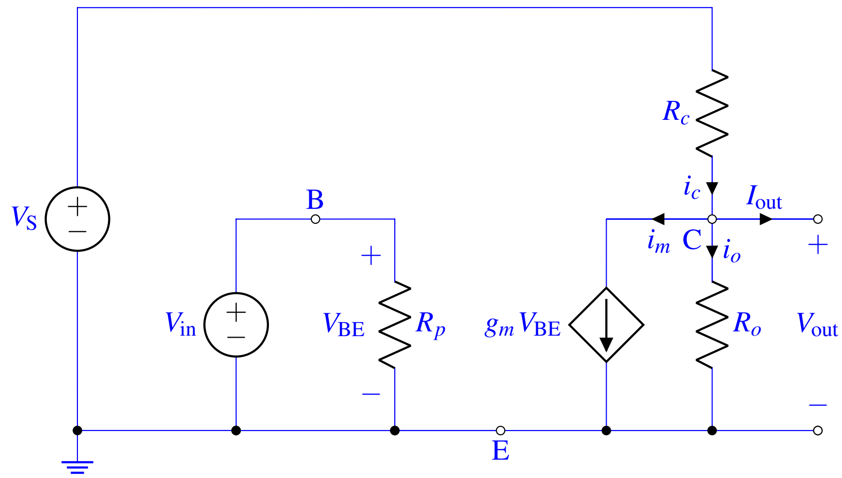


(b) Use the open circuit test to find the Thevenin voltage, V_{th}, between nodes C and E.

Recall the open circuit test finds $V_{\text{out}} = V_{\text{oc}}$ when an open circuit is connected across the terminals, then $V_{\text{th}} = V_{\text{oc}}$.

Solution:

Take the full equivalent circuit from part (a) and connect an open circuit across the output terminals.



Identify the KVL equation

$$V_{in} - V_{BE} = 0$$

which results in the dependent current source having a current of

$$g_m V_{BE} = g_m V_{in}$$

Next, identify a KCL equation at the output node and substitute the known node voltages

$$i_c - i_m - i_o - I_{out} = 0$$

$$\frac{V_S - V_{out}}{R_c} - g_m V_{in} - \frac{V_{out}}{R_o} - 0 = 0$$

Simplifying and rearranging this node voltage equation in terms of V_{out} is

$$V_{out} \cdot \left(\frac{1}{R_c} + \frac{1}{R_o}\right) = \frac{1}{R_c} V_S - g_m V_{in}$$

$$V_{out} = \left(\frac{1}{R_c} + \frac{1}{R_o}\right)^{-1} \cdot \left(\frac{1}{R_c} V_S - g_m V_{in}\right)$$

$$V_{out} = R_c || R_o \cdot \left(\frac{1}{R_c} V_S - g_m V_{in}\right)$$

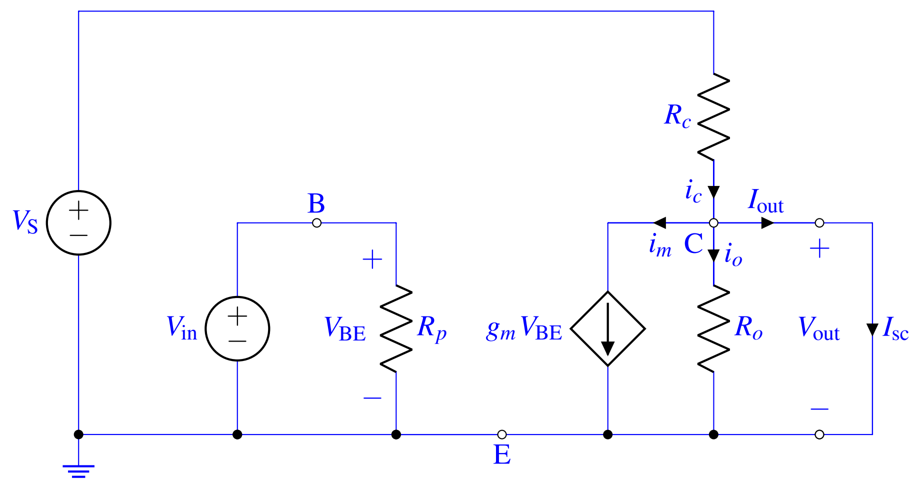
This concludes the open circuit test where the Thevenin voltage is

$$V_{th} = V_{oc} = V_{out} = R_c || R_o \cdot \left(\frac{1}{R_c} V_S - g_m V_{in}\right)$$

(c) Use the short circuit test to find the Norton current, I_{no}, between nodes C and E.
Recall the short circuit test finds $I_{out} = I_{sc}$ when a short circuit is connected between the terminals, then $I_{no} = I_{sc}$.

Solution:

Now for the short circuit test, connect a short circuit across the output terminals.



Similar to part (a), we identify the KVL equation

$$V_{in} - V_{BE} = 0$$

which results in the dependent current source having a current of

$$g_m V_{BE} = g_m V_{in}$$

Next, identify a KCL equation at the output node and substitute the known node voltages

$$i_c - i_m - i_o - I_{out} = 0$$
$$\frac{V_S - V_{out}}{R_c} - g_m V_{in} - \frac{V_{out}}{R_o} - I_{out} = 0$$

In this circuit, $V_{out} = 0$. Simplify and rearrange this node voltage equation in terms of I_{out} as

$$I_{out} = \frac{1}{R_c} V_S - g_m V_{in}$$

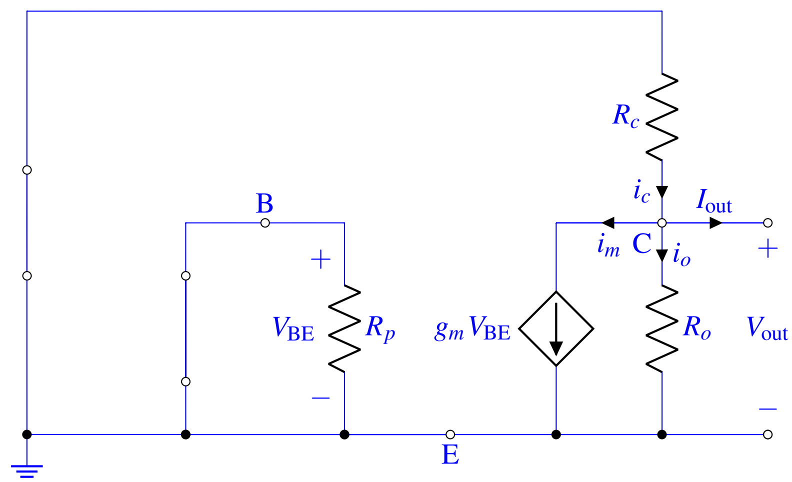
This concludes the short circuit test where the Norton current is

$$I_{no} = I_{sc} = \frac{1}{R_c} V_S - g_m V_{in}$$

(d) Find the Thevenin/Norton resistance $R_{th} = R_{no}$ using $R_{th} = \frac{V_{th}}{I_{no}}$.
Solution:

$$R_{th} = \frac{V_{th}}{I_{no}} = \frac{R_c || R_o \cdot \left(\frac{1}{R_c} V_S - g_m V_{in}\right)}{\left(\frac{1}{R_c} V_S - g_m V_{in}\right)}$$
$$= R_c || R_o$$

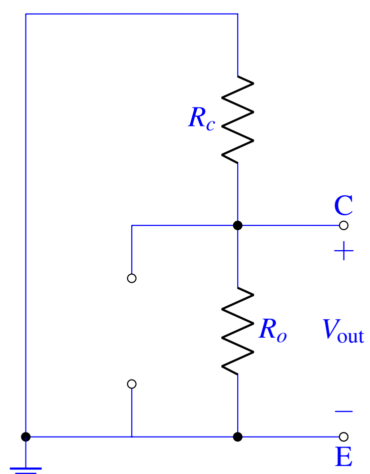
(e) We can also find R_{th} by turning off all of the independent sources (but *not* the dependent sources) and deriving the equivalent resistance seen from the terminals. Derive R_{th} with this method. Does it match your answer from part (c)?
Hint: To simplify the dependent source, focus on first finding V_{BE}.
Solution:

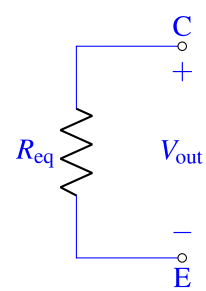


If we turn off the two independent voltage sources ($V = 0$), then

$$V_{\text{BE}} = 0 \quad \longrightarrow \quad g_m V_{\text{BE}} = 0$$

which results in the dependent current source being an equivalent open circuit ($I = 0$).





The final equivalent resistance R_{eq} is comprised of R_c in parallel with R_o.

$$R_{\text{th}} = R_{\text{eq}} = R_c || R_o$$

This Thevenin resistance R_{th} matches the value derived in part (c).

3. 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:

$$\langle \vec{x}, \vec{y} \rangle \;=\; x_1y_1 + x_2y_2 + \ldots + x_ny_n \;= \vec{x}^\mathsf{T}\vec{y}\,, \quad \text{for any } \; \vec{x}, \vec{y} \in \mathbb{R}^n$$

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

(a) $\langle \vec{x}, \vec{x} \rangle = \|\vec{x}\|^2$

Solution:

$$\begin{aligned}
\left\langle \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right\rangle &= x_1 \cdot x_1 + x_2 \cdot x_2 + \cdots + x_n \cdot x_n \\
&= x_1^2 + x_2^2 + \cdots + x_n^2 \\
&= (\sqrt{x_1^2 + x_2^2 + \cdots + x_n^2})^2 \\
&= (\|\vec{x}\|)^2
\end{aligned}$$

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

(b) $\langle -\vec{x}, \vec{y} \rangle = -\langle \vec{x}, \vec{y} \rangle$.

Solution:

$$\begin{aligned}
\langle -\vec{x}, \vec{y} \rangle &= \left\langle \begin{bmatrix} -x_1 \\ -x_2 \\ \vdots \\ -x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \right\rangle \\
&= -x_1 \cdot y_1 - x_2 \cdot y_2 - \cdots - x_n \cdot y_n \\
&= -(x_1 \cdot y_1 + x_2 \cdot y_2 + \cdots + x_n \cdot y_n) \\
&= -\langle \vec{x}, \vec{y} \rangle
\end{aligned}$$

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.

(c) $\langle \vec{x}, \vec{y} + \vec{z} \rangle = \langle \vec{x}, \vec{y} \rangle + \langle \vec{x}, \vec{z} \rangle$

Solution:

$$\begin{aligned}
\langle \vec{x}, \vec{y} + \vec{z} \rangle &= \vec{x}^\mathsf{T}(\vec{y} + \vec{z}) \\
&= \vec{x}^\mathsf{T}\vec{y} + \vec{x}^\mathsf{T}\vec{z} \\
&= \langle \vec{x}, \vec{y} \rangle + \langle \vec{x}, \vec{z} \rangle
\end{aligned}$$

The inner product is distributive.

(d) $\langle \vec{x}+\vec{y}, \vec{x}+\vec{y}\rangle = \langle \vec{x},\vec{x}\rangle + 2\langle \vec{x},\vec{y}\rangle + \langle \vec{y},\vec{y}\rangle$

Solution: Using the distributive and commutative properties from the previous parts,

$$\begin{aligned}
\langle \vec{x}+\vec{y}, \vec{x}+\vec{y}\rangle &= \langle \vec{x}, \vec{x}+\vec{y}\rangle + \langle \vec{y}, \vec{x}+\vec{y}\rangle \\
&= \langle \vec{x},\vec{x}\rangle + \langle \vec{x},\vec{y}\rangle + \langle \vec{y},\vec{x}\rangle + \langle \vec{y},\vec{y}\rangle \\
&= \langle \vec{x},\vec{x}\rangle + 2\langle \vec{x},\vec{y}\rangle + \langle \vec{y},\vec{y}\rangle
\end{aligned}$$

4. Inner Products

For each of the following functions, show whether it defines an inner product on the given vector space. If not, give a counterexample, i.e., find a pair of vectors p and q such that the given function fails to satisfy one of the inner product properties.

(a) For $\mathbb{R}^2$:

$$\langle \vec{p}, \vec{q}\rangle = \vec{p}^T \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \vec{q}$$

Solution: Yes, the function defines an inner product on $\mathbb{R}^2$. To show this, we will show that all of the three axioms apply. Let $\vec{p} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$ and $\vec{q} = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}$.

i. Symmetry:

$$\begin{aligned}
\langle \vec{p}, \vec{q}\rangle &= \begin{bmatrix} p_1 & p_2 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \\
&= \begin{bmatrix} p_1 & p_2 \end{bmatrix} \begin{bmatrix} 3q_1 + q_2 \\ q_1 + 2q_2 \end{bmatrix} \\
&= 3p_1q_1 + p_1q_2 + p_2q_1 + 2p_2q_2 \\
&= \begin{bmatrix} q_1 & q_2 \end{bmatrix} \begin{bmatrix} 3p_1 + p_2 \\ p_1 + 2p_2 \end{bmatrix} \\
&= \begin{bmatrix} q_1 & q_2 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \\
&= \langle \vec{q}, \vec{p}\rangle
\end{aligned}$$

ii. Linearity: Let $\vec{p}_1, \vec{p}_2, \vec{q} \in \mathbb{R}^2$.

$$\begin{aligned}
\langle \alpha\vec{p}_1 + \beta\vec{p}_2, q\rangle &= (\alpha\vec{p}_1 + \beta\vec{p}_2)^T \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \vec{q} \\
&= \alpha\vec{p}_1^{\,T} \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \vec{q} + \beta\vec{p}_2^{\,T} \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \vec{q} \\
&= \alpha\langle \vec{p}_1, \vec{q}\rangle + \beta\langle \vec{p}_2, \vec{q}\rangle
\end{aligned}$$

iii. Positive-definiteness:

$$\langle \vec{p}, \vec{p} \rangle = 3p_1p_1 + p_1p_2 + p_2p_1 + 2p_2p_2 = 3p_1^2 + 2p_2^2 + 2p_1p_2 = 2p_1^2 + p_2^2 + (p_1 + p_2)^2$$

Since all of the components are non-negative, $\langle \vec{p}, \vec{p} \rangle \geq 0$.
Furthermore, the inner product will be 0 if and only if $p_1 = p_2 = 0 \implies \vec{p} = \vec{0}$.

(b) For $\mathbb{R}^2$:

$$\langle \vec{p}, \vec{q} \rangle = \vec{p}^T \begin{bmatrix} 3 & 1 \\ 1 & -2 \end{bmatrix} \vec{q}$$

Solution:
No, the function does not define an inner product on $\mathbb{R}^2$. To show this, we will give a counterexample for the positive-definiteness axiom. Let $\vec{p} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$.

$$\begin{aligned}\langle \vec{p}, \vec{p} \rangle &= \begin{bmatrix} p_1 & p_2 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \\ &= 3p_1p_1 + p_1p_2 + p_2p_1 - 2p_2p_2 \\ &= 3p_1^2 - 2p_2^2 + 2p_1p_2\end{aligned}$$

Let $p_1 = 0$ and $p_2 = 1$. Then $3p_1^2 - 2p_2^2 + 2p_1p_2 = -2 < 0$.

(c) For $\mathbb{R}^2$:

$$\langle \vec{p}, \vec{q} \rangle = \vec{p}^T \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \vec{q}$$

Solution:
No, the function does not define an inner product on $\mathbb{R}^2$. To show this, we will give a counterexample for the symmetry axiom. Let $\vec{p} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$ and $\vec{q} = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}$.

$$\begin{aligned}\langle \vec{p}, \vec{q} \rangle &= \begin{bmatrix} p_1 & p_2 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \\ &= 3p_1q_1 + p_1q_2 + 2p_2q_2\end{aligned}$$

$$\begin{aligned}\langle \vec{q}, \vec{p} \rangle &= \begin{bmatrix} q_1 & q_2 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \\ &= 3p_1q_1 + p_2q_1 + 2p_2q_2\end{aligned}$$

Since there exists p_1, q_1, p_2, q_2 such that $p_1q_2 \neq p_2q_1$, we have $\langle \vec{p}, \vec{q} \rangle \neq \langle \vec{q}, \vec{p} \rangle$.

(d) For $\mathbb{R}^{2x2}$, the space of all 2x2 real matrices, the *Frobenius* inner product is defined as:

$$\langle A, B \rangle_F = Tr(A^T B)$$

Where A and B are 2x2 real matrices, and Tr represents the *trace* of a matrix, or the sum of its diagonal entries. Prove that the Frobenius inner product is valid over $\mathbb{R}^{2x2}$.

Solution:
To simplify analysis, let's define the entries of A and B as follows:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} A^T = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$$
$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} B^T = \begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{bmatrix}$$

We can now proceed to prove the three axioms.

i. Symmetry: Let's first compute the result of $A^T B$:

$$A^T B = \begin{bmatrix} a_{11}b_{11} + a_{21}b_{21} & a_{11}b_{12} + a_{21}b_{22} \\ a_{12}b_{11} + a_{22}b_{21} & a_{12}b_{12} + a_{22}b_{22} \end{bmatrix}$$

The trace (sum of diagonal elements) is:

$$a_{11}b_{11} + a_{21}b_{21} + a_{12}b_{12} + a_{22}b_{22}$$

Proceeding similarly, we get that the product of $B^T A$ is:

$$B^T A = \begin{bmatrix} a_{11}b_{11} + a_{21}b_{21} & b_{11}a_{12} + b_{21}a_{22} \\ b_{12}a_{11} + b_{22}a_{21} & a_{12}b_{12} + a_{22}b_{22} \end{bmatrix}$$

We can see that the diagonal elements, and hence the traces are equal, proving symmetry–

$$\langle A, B \rangle_F = \langle B, A \rangle_F$$

ii. Linearity: Let's define $A, B, C \in \mathbb{R}^{2x2}$ and $\alpha, \beta \in R$. Defining the elements in the same way as in the Symmetry section, lets look at

$$\begin{aligned} \langle \alpha A + \beta B, C \rangle_F &= Tr((\alpha A + \beta B)^T C) \\ Tr((\alpha A + \beta B)^T C) &= Tr(C^T(\alpha A + \beta B)) \\ &= Tr(C^T \alpha A + C^T \beta B) \\ &= Tr(\alpha C^T A) + Tr(\beta C^T B) \\ &= \alpha Tr(C^T A) + \beta Tr(C^T B) \\ &= \alpha \langle A, C \rangle_F + \beta \langle B, C \rangle_F \end{aligned}$$

Proving linearity.

iii. Positive Definiteness: If we define the elements of a matrix, $A \in \mathbb{R}^{2x2}$, in the same fashion as above, we find that $A^T A$ takes on the following form:

$$A^T A = \begin{bmatrix} a_{11}^2 + a_{21}^2 & a_{11}a_{12} + a_{21}a_{22} \\ a_{12}a_{11} + a_{22}a_{21} & a_{12}^2 + a_{22}^2 \end{bmatrix}$$

With the trace:

$$Tr(A^T A) = a_{11}^2 + a_{21}^2 + a_{12}^2 + a_{22}^2$$

Since $a_{11}, a_{22}, a_{12}, a_{21} \in \mathbb{R}$: $\langle A, A \rangle_F \geq 0$, and 0 if and only if the diagonal of A is 0.

5. Orthonormal Matrices

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

Orthornomal 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.

Hint: The transpose of a product of matrices is equivalent to the product of the transposes of the matrices in reverse order. For example $(U\vec{x})^T = \vec{x}^T U^T$.

(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:

$$\langle U\vec{x},U\vec{y}\rangle = (U\vec{x})^T(U\vec{y}) = \vec{x}^T U^T U\vec{y} = \vec{x}^T U^{-1}U\vec{y} = \vec{x}^T\vec{y} = \langle \vec{x},\vec{y}\rangle$$

(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 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.

6. Audio File Matching

Learning Goal: *This problem motivates the application of correlation for pattern matching applications such as Shazam. Note: Shazam is an application that identifies songs playing around you.*

Many audio processing applications rely on representing audio files as vectors, referred to as audio *signals*. Every component of the vector determines the sound we hear at a given time. We can use inner products to determine if a particular audio clip is part of a longer song, similar to an application like *Shazam*.

Let us consider a very simplified model for an audio signal, $\vec{x}$. At each timestep k, the audio signal can be either $x[k] = -1$ or $x[k] = 1$.

(a) Say we want to compare two audio files of the same length N to decide how similar they are. First, consider two vectors that are exactly identical, namely $\vec{x}_1 = \begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}^T$ and $\vec{x}_2 = \begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}^T$. What is the inner product of these two vectors? What if $\vec{x}_1 = \begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}^T$ but $\vec{x}_2$ oscillates between 1 and -1 (i.e., $\vec{x}_1 = \begin{bmatrix}1 & -1 & 1\cdots & -1\end{bmatrix}^T$)? Assume that N, the length of the two vectors, is an even number.
Use this to suggest a method for comparing the similarity between a generic pair of length-N vectors.
Solution:

The inner product of $\vec{x}_1 = \begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}^T$ and $\vec{x}_2 = \begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}^T$ is $\vec{x}_1 \cdot \vec{x}_2 = N$. The inner product of $\vec{x}_1 = \begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}^T$ and $\vec{x}_2 = \begin{bmatrix}1 & -1 & 1 & -1 & \cdots & 1 & -1\end{bmatrix}^T$ is $\vec{x}_1 \cdot \vec{x}_2 = 0$ when the vector length is even. To see this, take the sum of the first two terms of each vector.

$$\vec{x}_{1,1} \cdot \vec{x}_{2,1} + \vec{x}_{1,2} \cdot \vec{x}_{2,2} = 1 \cdot 1 + 1 \cdot -1 = 0$$

This yields zero, and repeats multiple times, leading to a total sum of 0. To compare two vectors of length N composed of 1 and -1, we take the inner product of the two vectors, a large inner product means the vectors have a similar direction.

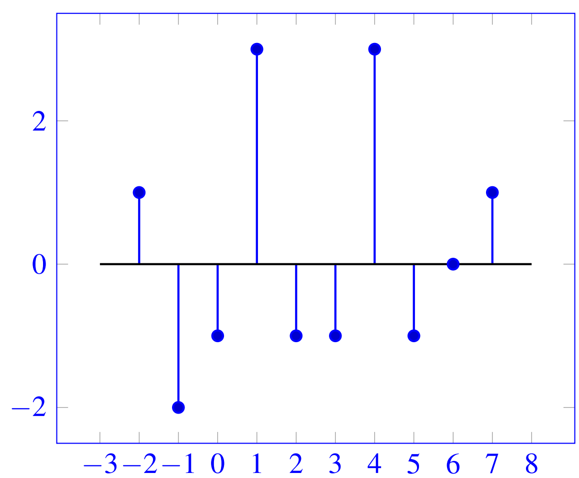
In many circumstances, an inner product with a very large negative value would mean the vectors are very different, but it turns out that humans are unable to perceive the sign of sound, so two sounds vectors $\vec{x}$ and $-\vec{x}$ sound exactly the same. As a result, for this problem we are interested in is the **absolute value** of the dot product, but in many other problems, we will interpret a large negative dot product as very different vectors. Don't take off points in parts (a), (b), or (c) if you didn't mention the absolute value.

(b) Next, suppose we want to find a short audio clip in a longer one. We might want to do this for an application like *Shazam*, which is able to identify a song from a short clip. Consider the vector of length 8, $\vec{x} = \begin{bmatrix}-1 & 1 & 1 & -1 & 1 & 1 & -1 & 1\end{bmatrix}^T$.

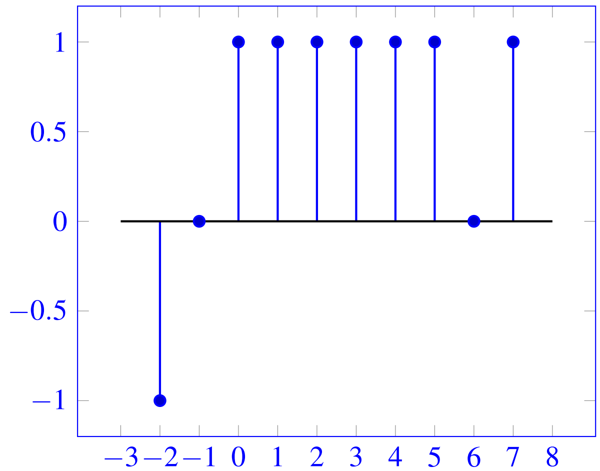
We want to find the short segment $\vec{y} := \begin{bmatrix}y[0] & y[1] & y[2]\end{bmatrix}^T = \begin{bmatrix}1 & 1 & -1\end{bmatrix}^T$ in the longer vector. To do this, perform the linear cross correlation between these two finite length sequences and identify at what shift(s) the linear cross correlation is maximized. Apply the same technique to identify what shift(s) gives the best match for $\vec{y} = \begin{bmatrix}1 & 1 & 1\end{bmatrix}^T$.

(If you wish, you may use iPython to do this part of the question, but you do not have to.)

Solution:



The above plot is $\text{corr}_{\vec{x}}(\vec{y})[k]$ where $\vec{y} = \begin{bmatrix}1 & 1 & -1\end{bmatrix}^T$. At shifts 1 and 4 the cross correlation is its *maximum possible value*, 3. These are both good matches.



The above plot is $\text{corr}_{\vec{x}}(\vec{y})[k]$ where $\vec{y} = \begin{bmatrix}1 & 1 & 1\end{bmatrix}^T$. At shifts 0 through 5 the cross correlation is only 1. There is not a really good match like before.

(c) Now suppose our audio vector is represented using integers beyond simply just 1 and -1. Find the short audio clip $\vec{y} = \begin{bmatrix}1 & 2 & 3\end{bmatrix}^T$ in the song given by $\vec{x} = \begin{bmatrix}1 & 2 & 3 & 1 & 2 & 2 & 3 & 10\end{bmatrix}^T$. Where do you expect to see the peak in the correlation of the two signals? Is the peak where you want it to be, i.e. does it pull out the clip of the song that you intended? Why?

(If you wish, you may use iPython to do this part of the question, but you do not have to.)

Solution:

Applying the technique in part (b), we get the best match to be $\begin{bmatrix}2 & 3 & 10\end{bmatrix}^T$ as this has the largest dot product with $\vec{y} = \begin{bmatrix}1 & 2 & 3\end{bmatrix}^T$. This is not where we expect to see the peak, as we observe the short audio clip $\vec{y}$ appears at the beginning of the song.

This happens because the volume at the end of the song is louder than the beginning of the song. Despite the angle not matching as well, the louder volume causes the linear cross correlation to be larger.

(d) Let us think about how to get around the issue in the previous part. We applied cross-correlation to compare segments of $\vec{x}$ of length 3 (which is the length of $\vec{y}$) with $\vec{y}$. Instead of directly taking the cross correlation, we want to normalize each inner product computed at each shift by the magnitudes of both segments, i.e. we want to consider the inner product $< \frac{\vec{x}_k}{\|\vec{x}_k\|}, \frac{\vec{y}}{\|\vec{y}\|} >$, where $\vec{x}_k$ is the length 3 segment starting from the k-th index of $\vec{x}$. This is referred to as normalized cross correlation. Using this procedure, now which segment matches the short audio clip best?

Solution: Using the normalized cross correlation procedure, the maximum inner product occurs when the two vectors are the same. This is because every vector now has unit magnitude (due to the normalization) and so the inner product only depends on the difference in the angles between the vectors. In our case, the 0^{th} shift of the autocorrelation will produce the largest magnitude since the two vectors $\frac{\vec{y}}{\|\vec{y}\|} = \begin{bmatrix}1 & 2 & 3\end{bmatrix}^T / \sqrt{14}$ and $< \frac{\vec{x}_0}{\|\vec{x}_0\|} = \begin{bmatrix}1 & 2 & 3\end{bmatrix}^T / \sqrt{14}$ are the same. This matches our desired behavior!

(e) We can use this on a more ‘realistic’ audio signal – refer to the IPython notebook, where we use normalized cross-correlation on a real song. Run the cells to listen to the song we are searching through, and add a simple comparison function `vector_compare` to find where in the song the clip comes from (i.e. write down the matching timestamp of the long audio clip). Running this may take a couple minutes on your machine, but note that this computation can be highly optimized and run super fast in the real world! Also note that this is not exactly how Shazam works, but it draws heavily on some of these basic ideas.

Note: if the script is running super slowly on Datahub, we recommend running it locally by installing Jupyter Notebook. An explanation for how to do this can be found here.

Solution:

See `sol13.ipynb`.

7. 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. $\|e\|^2 = \|\text{proj}_{\vec{a}}(\vec{b}) - \vec{b}\|^2$?

Solution:

$$\text{proj}_{\vec{a}}(\vec{b}) = \frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{a}\|^2}\vec{a} = \frac{\vec{a}^T\vec{b}}{\|\vec{a}\|^2}\vec{a} \tag{1}$$

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

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

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

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

$\|e\|^2 = \| \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix} \|^2$

$\|e\|^2 = \| \begin{bmatrix} -2 \\ -2 \\ 2 \end{bmatrix} \|^2$

$\|e\|^2 = 4 + 4 + 4 = 12$.

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

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

We will compute $\hat{\vec{x}}$ by solving the following least squares problem,

$$\min_{\vec{x}} \|\mathbf{A}\vec{x} - \vec{b}\|^2 \tag{2}$$

The solution yields,

$$\begin{aligned}
\hat{\vec{x}} &= (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\vec{b} &(3)\\
&= \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}^T \begin{bmatrix} 1 \\ 4 \\ -5 \end{bmatrix} &(4)\\
&= \begin{bmatrix} 0.5 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -4 \\ 4 \end{bmatrix} &(5)\\
&= \begin{bmatrix} -2 \\ 4 \end{bmatrix} &(6)
\end{aligned}$$

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

The squared error between the projection and $\vec{b}$ is $\|\vec{e}\|^2 = \left\| \begin{bmatrix} -2 \\ 4 \\ -2 \end{bmatrix} - \begin{bmatrix} 1 \\ 4 \\ -5 \end{bmatrix} \right\|^2 = 18.$

8. Mechanical Trilateration

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

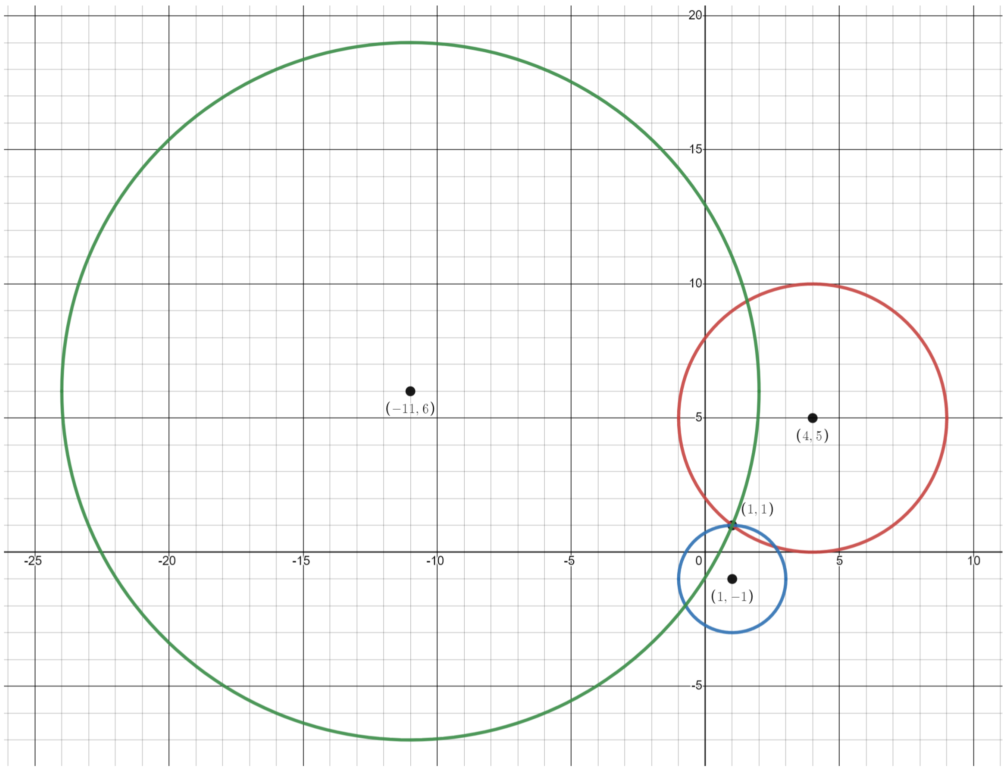
Trilateration is the problem of finding one's coordinates given distances from known beacon locations. For each of the following trilateration problems, you are given 3 beacon locations ($\vec{s}_1$, $\vec{s}_2$, $\vec{s}_3$) and the corresponding distance (d_1, d_2, d_3) from each beacon to your location.

For each problem, **graph** (by hand, with a graphing calculator, or iPython) the set of coordinates indicating your possible location for each beacon and find any coordinate solutions where they all intersect. Then **solve the trilateration problem algebraically** using the method introduced in lecture, to find your location or possible locations. If a solution does not exist, state that it does not.

For any solutions found using trilateration, be sure to check that they are consistent with the beacon measurements.

(a) $\vec{s}_1 = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$, $d_1 = 5$; $\vec{s}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $d_2 = 2$; $\vec{s}_3 = \begin{bmatrix} -11 \\ 6 \end{bmatrix}$, $d_3 = 13$

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.

$$\|\vec{x}-\vec{s}_1\|^2 = d_1^2$$
$$\|\vec{x}-\vec{s}_2\|^2 = d_2^2$$
$$\|\vec{x}-\vec{s}_3\|^2 = d_3^2$$

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

$$\|\vec{x}-\vec{s}_i\|^2 = \langle \vec{x}-\vec{s}_i, \vec{x}-\vec{s}_i\rangle = (\vec{x}-\vec{s}_i)^T(\vec{x}-\vec{s}_i)$$

$$\vec{x}^T\vec{x} - 2\vec{x}^T\vec{s}_1 + \vec{s}_1^T\vec{s}_1 = d_1^2$$
$$\vec{x}^T\vec{x} - 2\vec{x}^T\vec{s}_2 + \vec{s}_2^T\vec{s}_2 = d_2^2$$
$$\vec{x}^T\vec{x} - 2\vec{x}^T\vec{s}_3 + \vec{s}_3^T\vec{s}_3 = d_3^2$$

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

$$2\vec{x}^T\vec{s}_3 - 2\vec{x}^T\vec{s}_1 = d_1^2 - d_3^2 + \vec{s}_3^T\vec{s}_3 - \vec{s}_1^T\vec{s}_1$$
$$2\vec{x}^T\vec{s}_3 - 2\vec{x}^T\vec{s}_2 = d_2^2 - d_3^2 + \vec{s}_3^T\vec{s}_3 - \vec{s}_2^T\vec{s}_2$$

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

$$\begin{bmatrix} 2(\vec{s}_3-\vec{s}_1)^T \\ 2(\vec{s}_3-\vec{s}_2)^T \end{bmatrix}\vec{x} = \begin{bmatrix} d_1^2 - d_3^2 + \|\vec{s}_3\|^2 - \|\vec{s}_1\|^2 \\ d_2^2 - d_3^2 + \|\vec{s}_3\|^2 - \|\vec{s}_2\|^2 \end{bmatrix}$$

We have that:

$$2(\vec{s}_3 - \vec{s}_1) = \begin{bmatrix} -30 \\ 2 \end{bmatrix}$$
$$2(\vec{s}_3 - \vec{s}_2) = \begin{bmatrix} -24 \\ 14 \end{bmatrix}$$
$$d_1^2 - d_3^2 + \|\vec{s}_3\|^2 - \|\vec{s}_1\|^2 = 25 - 169 + 157 - 41 = -28$$
$$d_2^2 - d_3^2 + \|\vec{s}_3\|^2 - \|\vec{s}_2\|^2 = 4 - 169 + 157 - 2 = -10$$

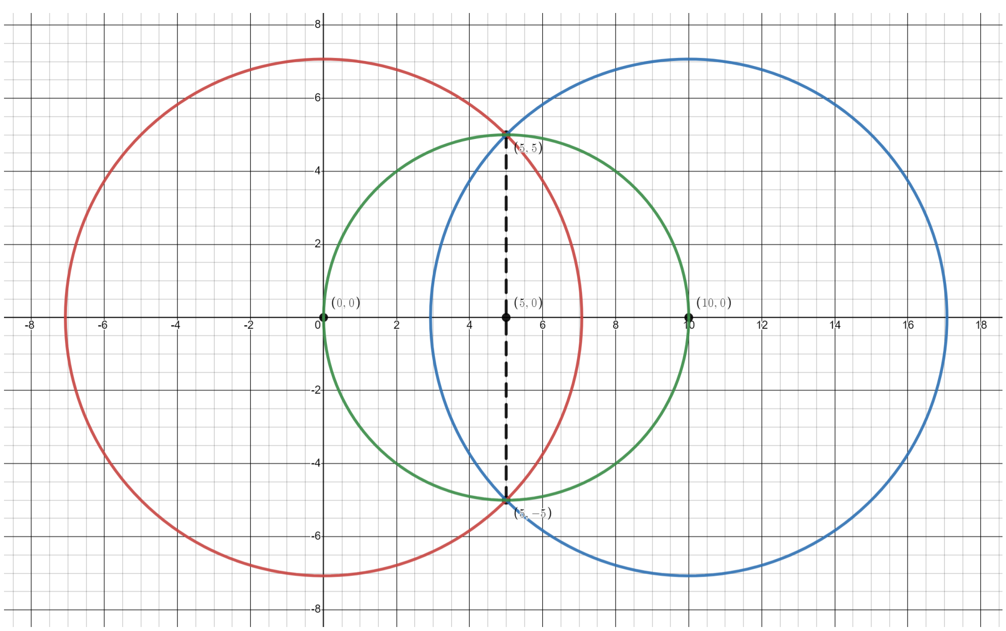
Which gives us the system $\begin{bmatrix} -30 & 2 \\ -24 & 14 \end{bmatrix} \vec{x} = \begin{bmatrix} -28 \\ -10 \end{bmatrix}$ with solution $\vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
A solution existing for this system of linear equations does not necessarily guarantee consistency of the system of nonlinear equations, but we can validate:

$$\left\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 4 \\ 5 \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} -3 \\ -4 \end{bmatrix} \right\|^2 = 25 = d_1^2$$
$$\left\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} 0 \\ 2 \end{bmatrix} \right\|^2 = 4 = d_2^2$$
$$\left\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} -11 \\ 6 \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} 12 \\ -5 \end{bmatrix} \right\|^2 = 169 = d_3^2$$

(b) $\vec{s}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, d_1 = 5\sqrt{2}; \vec{s}_2 = \begin{bmatrix} 10 \\ 0 \end{bmatrix}, d_2 = 5\sqrt{2}; \vec{s}_3 = \begin{bmatrix} 5 \\ 0 \end{bmatrix}, d_3 = 5$

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:

$$2(\vec{s}_3 - \vec{s}_1) = \begin{bmatrix} 10 \\ 0 \end{bmatrix}$$
$$2(\vec{s}_3 - \vec{s}_2) = \begin{bmatrix} -10 \\ 0 \end{bmatrix}$$
$$d_1^2 - d_3^2 + \|\vec{s}_3\|^2 - \|\vec{s}_1\|^2 = 50 - 25 + 25 - 0 = 50$$
$$d_2^2 - d_3^2 + \|\vec{s}_3\|^2 - \|\vec{s}_2\|^2 = 50 - 25 + 25 - 100 = -50$$

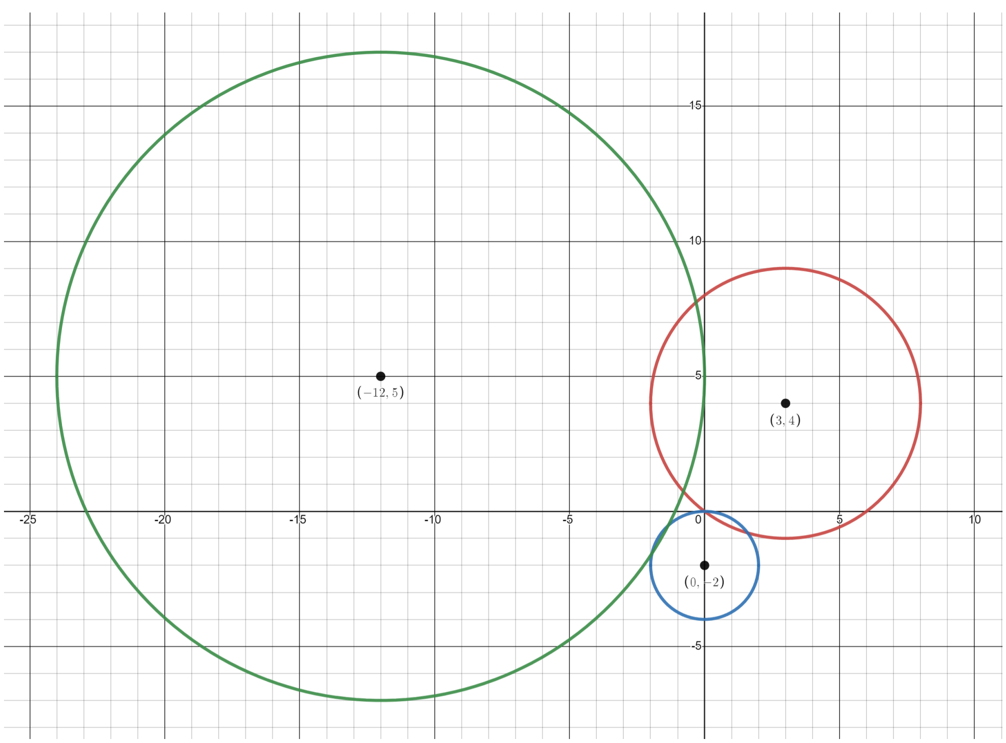
Which gives us the system $\begin{bmatrix} 10 & 0 \\ -10 & 0 \end{bmatrix} \vec{x} = \begin{bmatrix} 50 \\ -50 \end{bmatrix}$ with solution $\vec{x} = \begin{bmatrix} 5 \\ \alpha \end{bmatrix}$. We can see that by having collinear beacons, we may not be able to precisely determine our location (short exercise: how does this relate to span and vector spaces?)
However, from the graph we know that not all values of α are valid, so we can plug our solution back into the third distance equation:

$$\left\| \begin{bmatrix} 5 \\ \alpha \end{bmatrix} - \begin{bmatrix} 5 \\ 0 \end{bmatrix} \right\|^2 = 5^2 \implies \alpha^2 = 25 \implies \alpha = \pm 5$$

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 $\vec{x} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$ or $\vec{x} = \begin{bmatrix} 5 \\ -5 \end{bmatrix}$.

(c) $\vec{s}_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, d_1 = 5; \vec{s}_2 = \begin{bmatrix} 0 \\ -2 \end{bmatrix}, d_2 = 2; \vec{s}_3 = \begin{bmatrix} -12 \\ 5 \end{bmatrix}, d_3 = 12$

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:

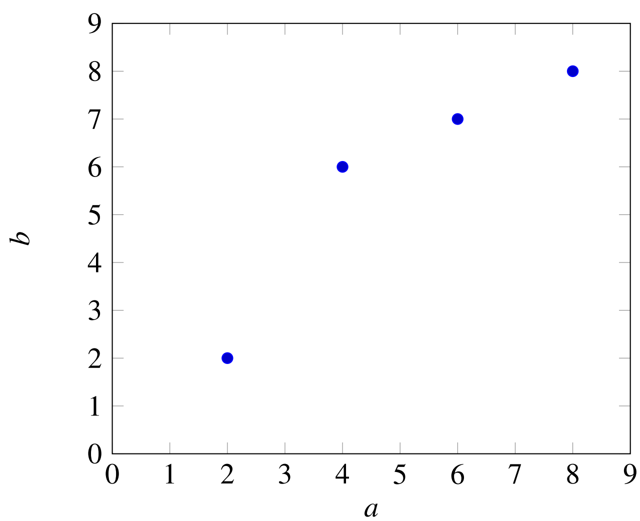
$$2(\vec{s}_3 - \vec{s}_1) = \begin{bmatrix} -30 \\ 2 \end{bmatrix}$$
$$2(\vec{s}_3 - \vec{s}_2) = \begin{bmatrix} -24 \\ 14 \end{bmatrix}$$
$$d_1^2 - d_3^2 + \|\vec{s}_3\|^2 - \|\vec{s}_1\|^2 = 25 - 144 + 169 - 25 = 25$$
$$d_2^2 - d_3^2 + \|\vec{s}_3\|^2 - \|\vec{s}_2\|^2 = 4 - 144 + 169 - 4 = 25$$

Which gives us the system $\begin{bmatrix} -30 & 2 \\ -24 & 14 \end{bmatrix} \vec{x} = \begin{bmatrix} 25 \\ 25 \end{bmatrix}$. While a solution, $\vec{x} = \begin{bmatrix} -\frac{75}{93} \\ \frac{75}{186} \end{bmatrix}$, for this system of linear equations exists, it will yield inconsistent distances when substituted back into the nonlinear equations.

$$\|\vec{s}_1 - x\|^2 = \left(3 + \frac{75}{93}\right)^2 + \left(4 - \frac{75}{186}\right)^2 = 27.43 \neq d_1^2 = 25$$

Therefore there is no solution.

9. Mechanical: Linear Least Squares



a	2	4	6	8
b	2	6	7	8

(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 $x = \hat{x}$ that minimizes the squared error

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

Note: By using this linear model, we are implicitly forcing the fit equation to go through the origin.

Do not use IPython for this calculation and show your work. Once you've computed $\hat{x}$, 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.)

Solution:

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

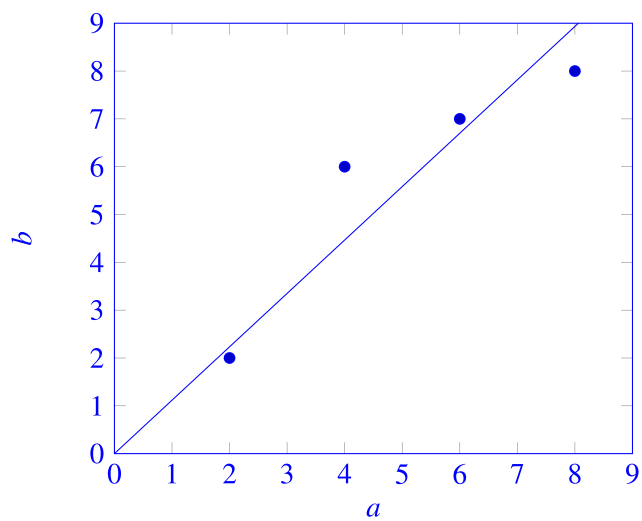
$$\begin{aligned} \hat{x} &= (\vec{a}^T\vec{a})^{-1}\vec{a}^T\vec{b} \\ &= \left(\begin{bmatrix} 2 & 4 & 6 & 8 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ 6 \\ 8 \end{bmatrix} \right)^{-1} \begin{bmatrix} 2 & 4 & 6 & 8 \end{bmatrix} \begin{bmatrix} 2 \\ 6 \\ 7 \\ 8 \end{bmatrix} \\ &= (120)^{-1}(134) = 1.1167 \end{aligned}$$

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

$$\begin{aligned} \vec{e} &= \vec{b} - \hat{\vec{b}} = \vec{b} - \hat{x}\vec{a} \\ &= \begin{bmatrix} 2 \\ 6 \\ 7 \\ 8 \end{bmatrix} - 1.1167 \begin{bmatrix} 2 \\ 4 \\ 6 \\ 8 \end{bmatrix} = \begin{bmatrix} -0.234 \\ 1.534 \\ 0.3 \\ -0.934 \end{bmatrix} \end{aligned}$$

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

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

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

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

Find x_1 and x_2 that minimize the squared error

$$\|\vec{e}\|^2 = \left\|\vec{b} - \mathbf{A}\vec{x}\right\|^2 = \left\| \begin{bmatrix} b_1 \\ \vdots \\ b_4 \end{bmatrix} - \begin{bmatrix} a_1 & 1 \\ \vdots & \vdots \\ a_4 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\|^2. \tag{8}$$

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

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

Solution:

Let $\vec{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$. Using the linear least squares formula with the new augmented $\mathbf{A}$ matrix, we calculate the optimal approximation of $\vec{x}$ as

$$\begin{aligned}
\hat{\vec{x}} &= \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\vec{b} \\
&= \left(\begin{bmatrix} 2 & 4 & 6 & 8 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 4 & 1 \\ 6 & 1 \\ 8 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 2 & 4 & 6 & 8 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 6 \\ 7 \\ 8 \end{bmatrix} \\
&= \begin{bmatrix} 120 & 20 \\ 20 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 4 & 6 & 8 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 6 \\ 7 \\ 8 \end{bmatrix} \\
&= \frac{1}{120(4) - 20(20)} \begin{bmatrix} 4 & -20 \\ -20 & 120 \end{bmatrix} \begin{bmatrix} 134 \\ 23 \end{bmatrix} \\
\hat{\vec{x}} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} &= \begin{bmatrix} 0.95 \\ 1 \end{bmatrix}
\end{aligned}$$

The linear model's prediction of $\vec{b}$ is given by

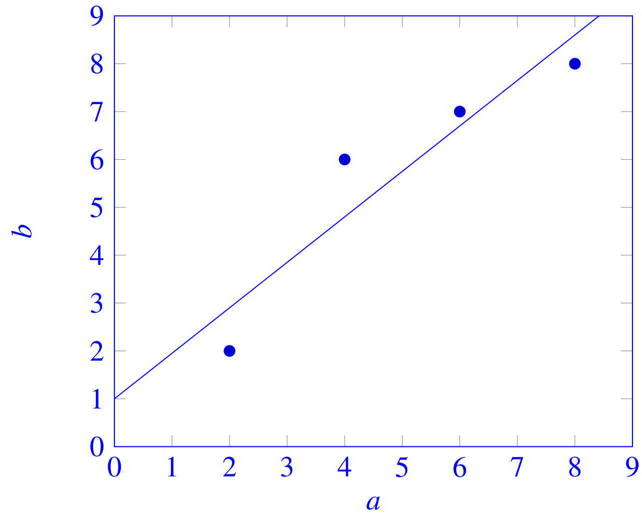
$$\hat{\vec{b}} = \mathbf{A}\hat{\vec{x}} = \begin{bmatrix} 2 & 1 \\ 4 & 1 \\ 6 & 1 \\ 8 & 1 \end{bmatrix} \begin{bmatrix} 0.95 \\ 1 \end{bmatrix} = \begin{bmatrix} 2.9 \\ 4.8 \\ 6.7 \\ 8.6 \end{bmatrix}$$

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

$$\|\vec{e}\|^2 = \vec{e}^T\vec{e} = 2.7$$



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.

10. Proof: Least Squares

Let $\vec{x}$ be the solution to a least squares problem. Show that the minimizing least squares error vector $\vec{e} = \vec{b} - \mathbf{A}\vec{x}$ is orthogonal to the columns of $\mathbf{A}$ by direct manipulation (i.e. plug the formula for the least squares solution $\vec{x}$ into the error vector and then check if $\mathbf{A}^T\vec{e} = \vec{0}$.)

Solution:

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

$$\begin{aligned}
\mathbf{A}^T\vec{e} = \mathbf{A}^T\left(\vec{b} - \mathbf{A}\vec{x}\right) &= \mathbf{A}^T\left(\vec{b} - \mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\vec{b}\right) \\
&= \mathbf{A}^T\vec{b} - \mathbf{A}^T\mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\vec{b} \\
&= \mathbf{A}^T\vec{b} - \mathbf{I}\mathbf{A}^T\vec{b} \\
&= \mathbf{A}^T\vec{b} - \mathbf{A}^T\vec{b} = \vec{0}
\end{aligned}$$

11. Trilateration With Noise!

Learning Goal: *This problem will help to understand how noise affects the accuracy of trilateration and consistency of the corresponding system of equations.*

In this question, we will explore how various types of noise affect the quality of triangulating a point on the 2D plane to see when trilateration works well and when it does not.

First, we will remind ourselves of the fundamental equations underlying trilateration.

(a) There are four beacons at the known coordinates $(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4)$. You are located at some unknown coordinate (x, y) that you want to determine. The distance between your location and each of the four beacons are d_1 through d_4, respectively. Write down one equation for each beacon that relates the coordinates to the distances using the Pythagorean Theorem.

Solution: For each beacon, we have the equation

$$(x-x_i)^2+(y-y_i)^2=d_i^2,$$

for $i\in\{1,2,3,4\}$.

(b) Unfortunately, the above system of equations is nonlinear, so we can't use least squares or Gaussian Elimination to solve it. We will use the technique discussed in lecture to obtain a system of linear equations. In particular, we can subtract the first of the above equations (involving x_1, y_1 and d_1) from the other three to obtain three linear equations (cancel out the nonlinear terms). Write down these three linear equations.

Combine the three equations in the above system into a single matrix equation of the form

$$\mathbf{A}\begin{bmatrix}x\\y\end{bmatrix}=\vec{b}.$$

Solution: Subtracting the 1st equation from the ith, we obtain

$$(x-x_i)^2-(x-x_1)^2+(y-y_i)^2-(y-y_1)^2=d_i^2-d_1^2,$$

so expanding and canceling the x^2 and y^2 terms, we obtain

$$-2xx_i+x_i^2+2xx_1-x_1^2-2yy_i+y_i^2+2yy_1-y_1^2=d_i^2-d_1^2$$

for $i\in\{2,3,4\}$.

Rearranging each of the above equations, we obtain

$$(-2x_i+2x_1)x+(-2y_i+2y_1)y=(d_i^2-x_i^2-y_i^2)-(d_1^2-x_1^2-y_1^2)$$

for $i\in\{2,3,4\}$. Stacking and writing in matrix form, we obtain

$$\begin{bmatrix}2(-x_2+x_1) & 2(-y_2+y_1)\\ 2(-x_3+x_1) & 2(-y_3+y_1)\\ 2(-x_4+x_1) & 2(-y_4+y_1)\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}(d_2^2-x_2^2-y_2^2)-(d_1^2-x_1^2-y_1^2)\\ (d_3^2-x_3^2-y_3^2)-(d_1^2-x_1^2-y_1^2)\\ (d_4^2-x_4^2-y_4^2)-(d_1^2-x_1^2-y_1^2)\end{bmatrix}.$$

(c) Now, go to the IPython notebook. In the notebook we are given three possible sets of measurements for the distances of each beacon from the receiver:

i. `ideal_distances`: the ideal set of measurements, the true distances of our receiver to the beacons. $d_1=d_2=d_3=d_4=5$.

ii. `imperfect_distances`: imperfect measurements. $d_1=5.5, d_2=4.5, d_3=5, d_4=5$.

iii. `one_bad_distances`: mostly perfect measurements, but d_1 is a very bad measurement. $d_1=7.5$ and $d_2=d_3=d_4=5$.

Plot the graph illustrating the case when the receiver has received `ideal_distances` and visually solve for the position of the observer (x,y). What is the coordinate?

Solution: From the plot, it is clear that $(x,y)=(0,0)$, since all four circles intersect at that point.

(d) You will now set up the above linear system using IPython. Fill in each element of the matrix $\mathbf{A}$ that you found in part (c).

Solution: $\mathbf{A}=\begin{bmatrix}2(-x_2+x_1) & 2(-y_2+y_1)\\ 2(-x_3+x_1) & 2(-y_3+y_1)\\ 2(-x_4+x_1) & 2(-y_4+y_1)\end{bmatrix}$.

See the IPython notebook for the actual code.

(e) Similarly, fill in the entries of $\vec{b}$ from part (c) in the `make_b` function.

Solution: $\vec{b} = \begin{bmatrix} (d_2^2 - x_2^2 - y_2^2) - (d_1^2 - x_1^2 - y_1^2) \\ (d_3^2 - x_3^2 - y_3^2) - (d_1^2 - x_1^2 - y_1^2) \\ (d_4^2 - x_4^2 - y_4^2) - (d_1^2 - x_1^2 - y_1^2) \end{bmatrix}$.

See the IPython notebook for the actual code.

(f) Now, you should be able to plot the estimated position of (x, y) using the supplied code for the `ideal_distances` observations. Modify the code to estimate (x, y) for `imperfect_distances` and `one_bad_distances`, and comment on the results.

In particular, for `one_bad_distances` would you intuitively have chosen the same point that our trilateration solution did knowing that only one measurement was bad?

Solution: We see that the solution to (x, y) moves away from the origin in the latter two cases. For `one_bad_distances`, even through three out of the four circles intersect at the origin (suggesting that $(x, y) = (0, 0)$), our least squares approach picks a point away from the origin, indicating that it might not be determining the best solution possible.

(g) We define the "cost" of a position (x, y) to be the sum of the squares of the differences in distance of that position from the observation, as defined symbolically in the notebook. Study the heatmap of the cost of various positions on the plane, and make sure you see why $(0, 0)$ appears to be the point with the lowest cost.

Now, compare the cost of $(0, 0)$ with the cost of your estimated position obtained from the least-squares solution in all three cases. For which cases does least squares do worse?

Solution: See IPython solutions. For `ideal_distances`, both approaches yield a cost of 0.0. For `imperfect_distances`, $(0, 0)$ is actually slightly worse than our least-squares solution, but in both cases the costs are fairly low.

For `one_bad_distances`, the costs are lower at $(0, 0)$ compared to the least-squares solution, as we expected from the heatmap.

12. Homework Process and Study Group

Who did you work with on this homework? List names and student ID's. (In case you met people at homework party or in office hours, you can also just describe the group.) How did you work on this homework? If you worked in your study group, explain what role each student played for the meetings this week.

Solution:

I first worked by myself for 2 hours, but got stuck on problem 5. Then I met with my study group.

XYZ played the role of facilitator ... etc. We were still stuck on problem 5 so we went to office hours to talk about the problem.

Then I went to homework party for a few hours, where I finished the homework.