CSM: EECS 16A (Designing Information Devices and Systems I)

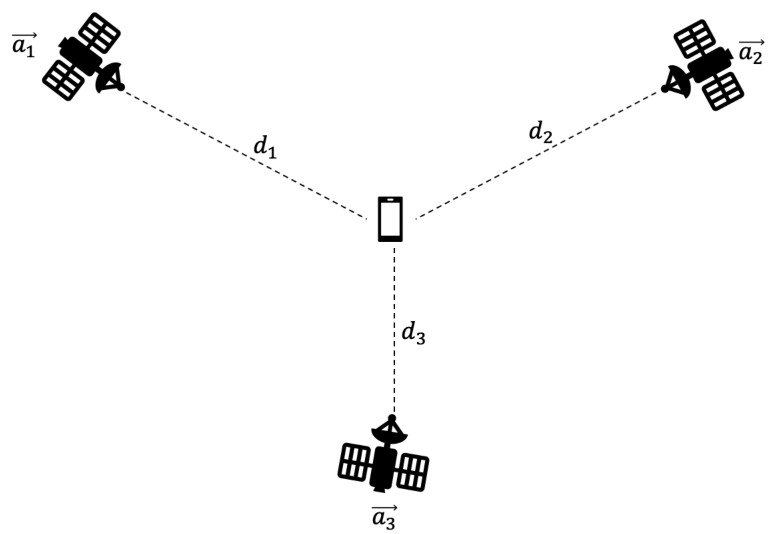
Worksheet #12

Term: **Fall 2019** *Name:*

Problem 1: GPS - Global Positioning System

Learning Goal: Understanding how trilateration works and how it is derived, working with norms and inner products, and applying cross correlation to find delays in received signals.

Suppose that you are the engineer tasked with the job of making Google Maps. For this, you want to be able to determine the position of a user using satellite information. In particular, assume that you know d_1, d_2 and d_3, the distances from the user's cellphone to 3 satellites. You know the positions of these satellites to be $\vec{a_1}, \vec{a_2}$, and $\vec{a_3}$. Here's a simplified figure demonstrating what's been given so far:



Note: What does it mean when we say "position"? You can assume that these positions are taken relative to some common origin. Say, the Google HQ - Mountain View, CA.

1. Suppose the user's location (or the phone's location) is given by the vector $\vec{x}$, write out a system of equations representing the distances from the user to all 3 satellites (Express your answer in terms of $\vec{x}, \vec{a_1}, \vec{a_2}, \vec{a_3}, d_1, d_2$, and d_3).

   Solution:

   $$\|\vec{x} - \vec{a_1}\| = d_1$$
   $$\|\vec{x} - \vec{a_2}\| = d_2$$
   $$\|\vec{x} - \vec{a_3}\| = d_3$$

   These equations relate the position of the user to the distance between the user and each satellite.

2. Rewrite these equations in terms of inner products of vectors. Are these equations linear with respect to $\vec{x}$?

Solution: By squaring each side, we get the following:

$$(\vec{x}-\vec{a_1})^T(\vec{x}-\vec{a_1}) = d_1^2$$
$$(\vec{x}-\vec{a_2})^T(\vec{x}-\vec{a_2}) = d_2^2$$
$$(\vec{x}-\vec{a_3})^T(\vec{x}-\vec{a_3}) = d_3^2$$

Then, expanding the LHS, we get:

$$\vec{x}^T\vec{x} - 2\vec{a_1}^T\vec{x} + \vec{a_1}^T\vec{a_1} = d_1^2$$
$$\vec{x}^T\vec{x} - 2\vec{a_2}^T\vec{x} + \vec{a_2}^T\vec{a_2} = d_2^2$$
$$\vec{x}^T\vec{x} - 2\vec{a_3}^T\vec{x} + \vec{a_3}^T\vec{a_3} = d_3^2$$

These equations are not linear in $\vec{x}$, because they contain an $\vec{x}^T\vec{x}$ term.

3. Are there any non-linear terms in the equations from the previous part? Using **elimination of variables**, rewrite everything as a system of **linear** equations.

   Solution: The non-linear terms in the equations are $\vec{x}^T\vec{x}$, but we can use variable elimination to get rid of them.

   Subtract the first equation from the other 2 in order to eliminate the $\vec{x}^T\vec{x}$ term. We get the following:

$$2(\vec{a_1}^T\vec{x} - \vec{a_2}^T\vec{x}) + (\vec{a_2}^T\vec{a_2} - \vec{a_1}^T\vec{a_1}) = d_2^2 - d_1^2$$
$$2(\vec{a_1}^T\vec{x} - \vec{a_3}^T\vec{x}) + (\vec{a_3}^T\vec{a_3} - \vec{a_1}^T\vec{a_1}) = d_3^2 - d_1^2$$

   We can rewrite these as a linear equations:

$$2(\vec{a_1}-\vec{a_2})^T\vec{x} = \vec{a_1}^T\vec{a_1} - \vec{a_2}^T\vec{a_2} + d_2^2 - d_1^2$$
$$2(\vec{a_1}-\vec{a_3})^T\vec{x} = \vec{a_1}^T\vec{a_1} - \vec{a_3}^T\vec{a_3} + d_3^2 - d_1^2$$

   Or, in matrix-vector form,

$$2\begin{bmatrix}(\vec{a_1}-\vec{a_2})^T\\(\vec{a_1}-\vec{a_3})^T\end{bmatrix}\vec{x} = \begin{bmatrix}\|a_1\|^2 - \|a_2\|^2 + d_2^2 - d_1^2\\\|a_1\|^2 - \|a_2\|^2 + d_3^2 - d_1^2\end{bmatrix}$$

4. **Using the system of linear equations we have from the previous part**, if the location of the user (i.e. $\vec{x}$ is a 3-dimensional vector), do we have sufficient information to solve for $\vec{x}$? If not, then how many satellites do you need to locate the user?

   Solution: No, this is not sufficient. We have only 2 equations, but 3 variables (Equivalently, in a matrix-vector equation representation $A\vec{x} = \vec{b}$, this would correspond to a 2×3 matrix A, implying that not all the column vectors in A are linearly independent, which means A is not invertible). So, this system is underdetermined, meaning that we don't have sufficient information to know the exact location of the user. In other words, **there could be more than one possible location for the user!** If instead we had 4 satellites, then by subtracting one equation from all the rest, we would have 3 equations, and we could then solve the system.

5. Suppose now in more generalized terms, we want to not only triangulate the user's position, but also keep track of other information about the user to make more customized analysis. Given that the vector representing the user location now contains a total of n entries, what is the minimum number of satellites we need to find that vector?

   Solution: First of all, we know that in order to solve for all n unique entries in the user location vector, we need a total of n linear equations (linearly independent).

Now, based on the previous part, we can see that during the process of variable elimination, we lose a total of 1 equation to reduce the equations down to a linear system. This means we need to have 1 more equation than n equations we originally planned.

Therefore, we would need at least a total of $n + 1$ satellites.

In real life, we won't actually be given the distances from the user to the satellites, either. In other words, we also need to figure out how far away the satellites are from us! Fortunately, as we have already learned in class, **cross correlation** is something that might come in handy for us to figure out the distances. For all the remaining parts of this question, we will use what we have learned about **cross correlation** to figure out what the distances from the user to the satellites are.

6. To figure out how far away the satellites are from us, we can use our phone to receive radio signals from the satellites in the orbit. Once we have received the signals, we can then compare them with a reference signal on our phone to figure out the time it takes for the signal to reach us. Given our original reference signal:

$$\vec{s} = \begin{bmatrix} -1 & -1 & -1 & 1 & -1 \end{bmatrix}^T,$$

and the three signals we received, each having a period of 4 (we will only show one period of each signal):

$$\vec{r}_1 = \begin{bmatrix} -1 & -1 & -1 & 1 \end{bmatrix}^T$$

$$\vec{r}_2 = \begin{bmatrix} 1 & -1 & 1 & 1 \end{bmatrix}^T$$

$$\vec{r}_3 = \begin{bmatrix} 1 & 1 & -1 & 1 \end{bmatrix}^T$$

Find the cross correlations $\text{corr}_{\vec{r}_1}(\vec{s})$, $\text{corr}_{\vec{r}_2}(\vec{s})$, and $\text{corr}_{\vec{r}_3}(\vec{s})$ between $\vec{s}$ and all three received signals respectively, and plot them out below.

Solution: Using the formula for cross correlation:

$$\text{corr}_{\vec{r}}(\vec{s})[k] = \sum_{i=-\infty}^{\infty} \vec{r}[i]\vec{s}[i-k],$$

we can find the cross correlations between $\vec{s}$ and all three received signals to be:

$$\begin{aligned} \text{corr}_{\vec{r}_1}(\vec{s}) &= \begin{bmatrix} 1 & 0 & 1 & 0 & 4 & 1 & 0 & -1 \end{bmatrix}^T, \\ \text{corr}_{\vec{r}_2}(\vec{s}) &= \begin{bmatrix} -1 & 2 & -3 & 0 & 0 & -1 & -2 & -1 \end{bmatrix}^T,. \\ \text{corr}_{\vec{r}_3}(\vec{s}) &= \begin{bmatrix} -1 & 0 & 1 & -4 & 0 & -1 & 0 & -1 \end{bmatrix}^T \end{aligned}$$

The plots for the cross-correlated signals are as follows:

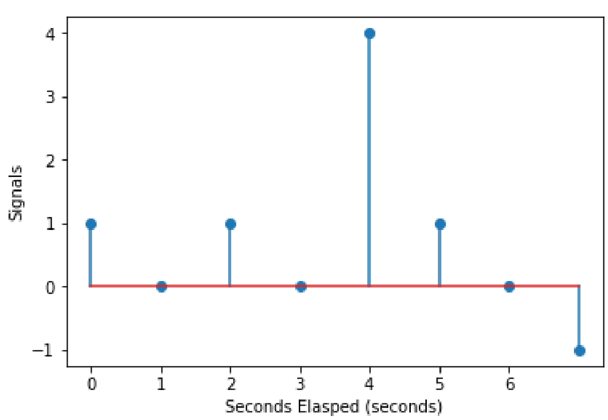


Figure 1: $\text{corr}_{\vec{r}_1}(\vec{s})$

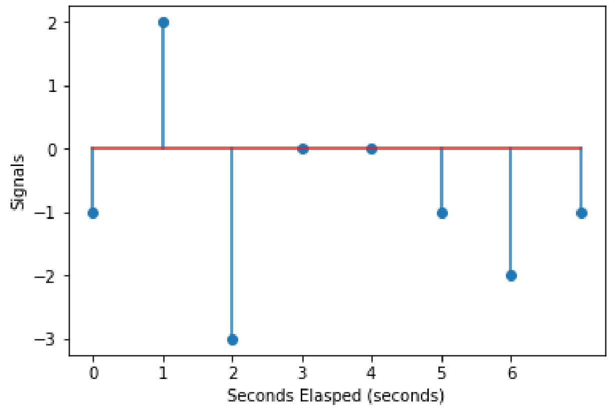


Figure 2: $\text{corr}_{\vec{r}_2}(\vec{s})$

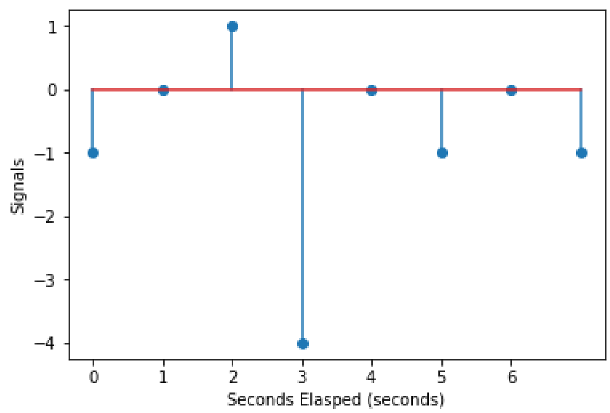


Figure 3: $\text{corr}_{\vec{r}_3}(\vec{s})$

7. Based on the cross-correlated signals, determine the delays (in seconds) for all 3 received signals $\vec{r}_1$, $\vec{r}_2$, and $\vec{r}_3$.

   Solution: Observing the plots for all 3 signals, we can see that the first received signal has a delay of 4 seconds, the second received signal has a delay of 1 second, and the last received signal has a delay of 2 seconds.

8. Given that the radio signal has a transmission speed of v, and assume all delays are relative to the source signal $\vec{s}$ (this means we assume $\vec{s}$ is received at time $t = 0$), find the distance d_1, d_2, and d_3 between the user location and the 3 satellites in orbit.

   Solution: Using the distance formula:

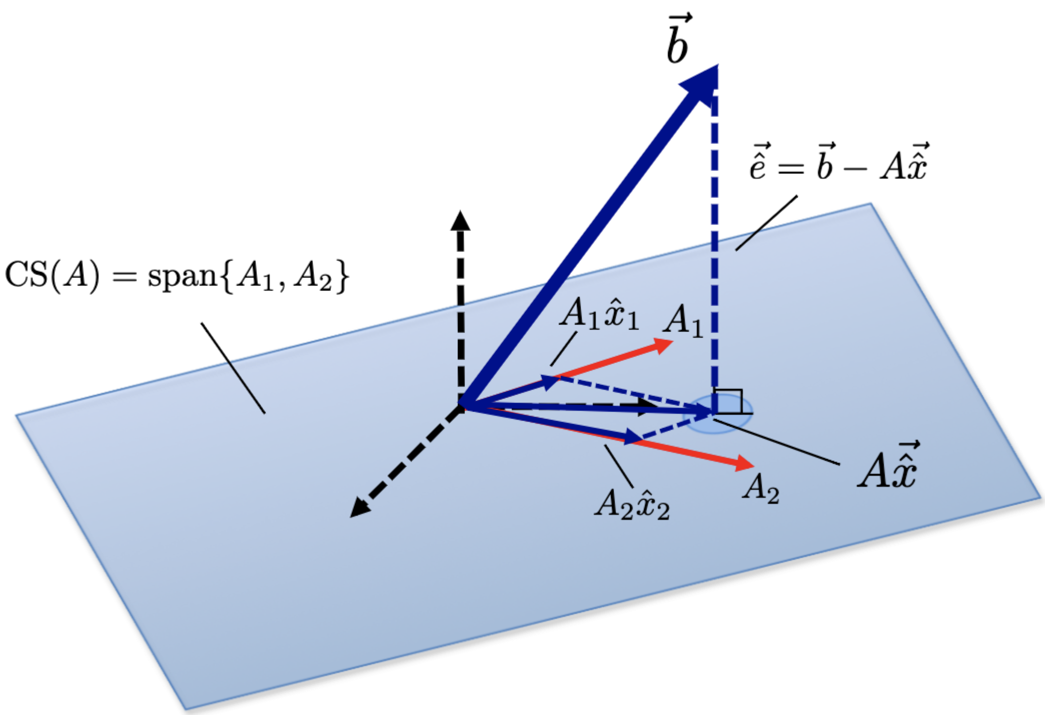
$$d = v \cdot \tau,$$

   where v is the transmission speed, τ is the delay (in seconds). we can find the distances between the user location and the satellites to be:

$$d_1 = 4v, \quad d_2 = v, \quad d_3 = 2v.$$

Problem 2: Least Squares - Geometric Intuitions

Learning Goal: Understanding of the mechanics and interpretations of squared norms, describing and deriving the least squares problem and solution, first geometrically, and then using calculus.



1. Consider that you have some equations of the form $\mathbf{A}\vec{x} = \vec{b}$, however, that there is no solution $\vec{x}$ that solves the equations. What does this tell us about $\vec{b}$ with respect to $\mathbf{A}_1, \mathbf{A}_2$ (the columns of $\mathbf{A}$)?

   Solution: $\vec{b}$ is not in the column space of $\mathbf{A}$. If there was some $\vec{x}$ which solved the equation $\mathbf{A}\vec{x} = \vec{b}$, then we could write $\vec{b} = \mathbf{A}_1 x_1 + \mathbf{A}_2 x_2$, and $\vec{b}$ would have been in the column space of $\mathbf{A}$. However, since there is no such $\vec{x}$, $\vec{b}$ is not in the column space of $\mathbf{A}$.

2. We know that there is no $\vec{x}$ that satisfies the equations exactly, but we still want to solve the equations to get a solution as close as possible.

   Let's say you had 3 choices, $\vec{x}_i, \vec{x}_j, \vec{x}_k$ (these are not drawn on the image). What could you compute in order to determine which of these would be the best choice instead of $\vec{x}$

   Solution: A good idea would be to multiply out $\mathbf{A}\vec{x}_i, \mathbf{A}\vec{x}_j, \mathbf{A}\vec{x}_k$, and see which one results in something closest to $\vec{b}$, and that's the one that we would pick.

3. Suppose that the real $\vec{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, and you have two close $\vec{x}_1, \vec{x}_2$, which result in possible $\vec{b}_1 = \begin{bmatrix} 1.5 \\ 1.5 \end{bmatrix}$ and another $\vec{b}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. How do we know which one is *closer*? What if we define *closer* to mean the sum of the components of the difference of the vectors.

   Solution: There are many ways to define closeness. One way is to simply sum up the differences in individual entries. For instance, the sum of the entries in $\vec{b} - \vec{b}_1$ vs. the sum of the entries in $\vec{b} - \vec{b}_2$. The first sum equals 0, while the second one equals 1. Huh? Both of these vectors seemed like they were '1' unit away, but the first one results in a sum of 0. This is because of negative numbers.

$\vec{b}_1$ is -0.5 away in the first element, and 0.5 away in the second. We should interpret this as being 1 away. In other words, we should take the absolute value of the differences of the individual elements. It turns out that this is, mathematically, annoying.

4. What is a different approach to solve the issue discussed above?

   Solution: Another way to get rid of negative numbers is by simply squaring the element-wise difference. Define the 'closeness' by the 'sum of the squares of individual elements in the vectors'.

5. Using this definition, how close is $\vec{b}_1$ to $\vec{b}$? How close is $\vec{b}_2$ to $\vec{b}$?

   Solution: $\vec{b}_1$: $-0.5^2 + 0.5^2 = 0.5$ and $\vec{b}_2$: $1^2 + 0^2 = 1$. Using this definition, $\vec{b}_1$ is closer, and that's the one we should pick!

6. More generally, we actually don't have a choice of just two or three $\vec{x}$s to pick to get as close to $\vec{b}$ as possible. We have an infinite number of choices. How can we tell which one is the best? Look at the image given at the top of the question, and decide something about $\vec{\hat{x}}$ and $\vec{\hat{e}}$.

   Solution: We want to pick an $\vec{\hat{x}}$ which results in the smallest norm-squared of $\vec{\hat{e}}$.

   $\vec{\hat{e}}$ is exactly the 'difference' that we discussed earlier of the correct $\vec{b}$ and the closest vector that we can get to $\vec{b}$. The norm-squared of $\vec{\hat{e}}$ is precisely the sum of squares of the elements in the difference. We want to minimize $\vec{\hat{e}}$.

7. Let's begin by recalling three facts. Recall that if we want to minimize some quantity squared, it is enough to minimize the quantity itself. Also, recall that given a point and a plane, the shortest line that one can possibly get starting at the point and ending at the plane, is a perpendicular line from the point to the plane. Finally, recall that if a vector $\vec{v}$ is orthogonal to another vector u, their inner product $\langle \vec{v}, \vec{u} \rangle = \vec{v}^T \vec{u} = 0$. Using this information, come up with a method to minimize the norm-squared of $\vec{\hat{e}}$.

   Solution: To minimize the norm-squared of $\vec{\hat{e}}$, we just want $\vec{\hat{e}}$ to be perpendicular to the column space of $\mathbf{A}$. This is the shortest-length (or smallest-norm) $\vec{\hat{e}}$ possible. Since finding the shortest-length vector is the same thing as minimizing the norm, which is equivalent to minimizing the squared norm.

   To make this vector orthogonal to the column space of $\mathbf{A}$, **it should be orthogonal to every column vector in A**. In other words, if

$$\mathbf{A} = \begin{bmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ \vec{a_1} & \vec{a_2} & \cdots & \vec{a_n} \\ \downarrow & \downarrow & \downarrow & \downarrow \end{bmatrix},$$

   using the definition of orthogonality between two vectors, we know that:

$$\begin{aligned} \left\langle \vec{a_1}, \mathbf{A}\vec{\hat{x}} - \vec{\hat{b}} \right\rangle &= \vec{a_1}^T(\mathbf{A}\vec{\hat{x}} - \vec{\hat{b}}) &= 0 \\ \left\langle \vec{a_2}, \mathbf{A}\vec{\hat{x}} - \vec{\hat{b}} \right\rangle &= \vec{a_2}^T(\mathbf{A}\vec{\hat{x}} - \vec{\hat{b}}) &= 0 \\ &\cdots \\ \left\langle \vec{a_n}, \mathbf{A}\vec{\hat{x}} - \vec{\hat{b}} \right\rangle &= \vec{a_n}^T(\mathbf{A}\vec{\hat{x}} - \vec{\hat{b}}) &= 0 \end{aligned}$$

   Now, note that:

$$\begin{bmatrix} \leftarrow & \vec{a_1}^T & \rightarrow \\ \leftarrow & \vec{a_2}^T & \rightarrow \\ & \vdots & \\ \leftarrow & \vec{a_n}^T & \rightarrow \end{bmatrix} = \mathbf{A}^T,$$

   more compactly, this allows us to write all the n equations above as: $\mathbf{A}^T(\mathbf{A}\vec{\hat{x}} - \vec{\hat{b}}) = \vec{0}$. This says $\mathbf{A}^T\mathbf{A}\vec{\hat{x}} = \mathbf{A}^T\vec{\hat{b}}$. The best $\vec{\hat{x}}$ we can find is $(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\vec{b}$.