EECS 16A Designing Information Devices and Systems I
Fall 2019 Lecture Notes Note 21

21.1 Module Goals

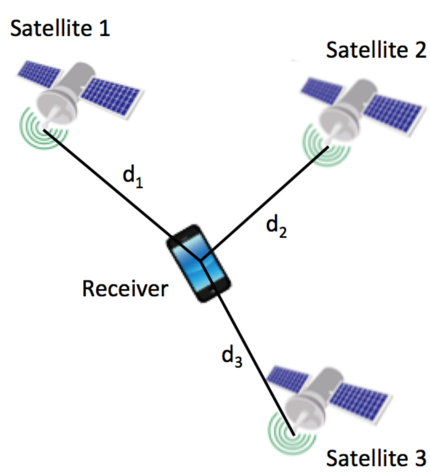
In this module, we introduce a family of ideas that are connected to **optimization** and **machine learning**, which are some of the modern pillars of EECS. You'll learn more about these concepts in EE 127 and EE 189.

This module uses the practical example of localization — figuring out where you are — as a source of examples and a connection to the lab. However, just like in previous modules, the ideas we will learn and the skills that we develop are far more generally applicable.

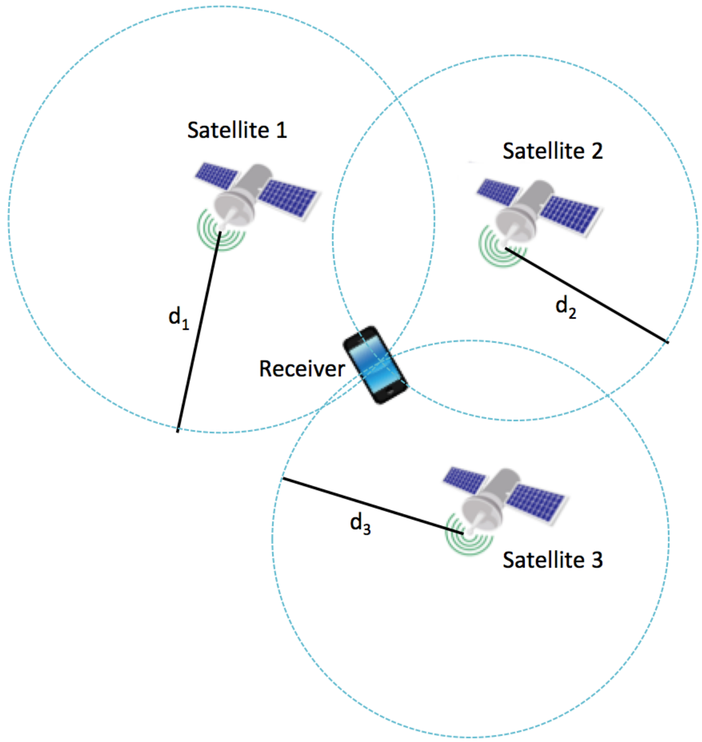
21.2 Introduction: Positioning Systems

GPS (the Global Positioning System) is a navigational system that we use all the time to tell us where we are. In this module, we'll explore the *signal processing* underlying GPS, and you'll build your own auditory positioning system in the lab.

How does GPS work? A receiver, such as your cell phone, receives messages from satellites orbiting the earth. From these messages, the receiver is able to determine how far away it is from each satellite. The receiver also knows the positions of each satellite.



The distance to the first satellite defines a set of possible locations where the receiver could be – a circle (or a sphere in 3D) around the satellite. Each other satellite also defines a circle of possible locations. By combining the measurements from all of the satellites, the receiver can determine it's position!



However, in real life, our measurements in will be noisy. Can we use extra measurements from more satellites to reduce the effects of noise? In this module, we'll also look at how to combine many measurements together to get a better estimate of position.

We can break this process into two parts, which we'll cover in detail in the upcoming notes:

1. Find the distances from the receiver to the satellites, based on the messages received.
2. Combine the measurements from each satellite to determine location.

In the upcoming notes, we'll explore all parts of this process in more detail. In this note, we'll introduce some fundamentals that will be built on in the rest of the module.

21.3 Inner Products

Definition of Inner Product: The (Euclidean) inner product between two vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$ is defined as:

$$
\begin{aligned}
\langle \vec{x}, \vec{y} \rangle \equiv \vec{x} \cdot \vec{y} \equiv \vec{x}^T \vec{y} & \quad (1) \\
= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} & \quad (2) \\
= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n & \quad (3) \\
= \sum_{i=1}^{n} x_i y_i & \quad (4)
\end{aligned}
$$

In physics, inner products are often called **dot products** (denoted by $\vec{x} \cdot \vec{y}$) and can be used to calculate values such as work. In this class we will use the notation $\langle \vec{x}, \vec{y} \rangle$ for the inner product.

Example 21.1 (Inner product of two vectors): Compute the inner product of the two vectors $\begin{bmatrix} -1 & 3.5 & 0 \end{bmatrix}^T$ and $\begin{bmatrix} 1 & 0 & 2 \end{bmatrix}^T$.

$$
\begin{aligned}
\left\langle \begin{bmatrix} -1 \\ 3.5 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \right\rangle = \begin{bmatrix} -1 & 3.5 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} & \quad (5) \\
= -1 \times 1 + 3.5 \times 0 + 0 \times 2 & \quad (6) \\
= -1 + 0 + 0 = -1. & \quad (7)
\end{aligned}
$$

21.4 Properties of Inner Products

Commutative: The inner product is a *commutative* or *symmetric* operation; that is, it remains the same even if we switch its arguments.[1]

Proof:

$$
\begin{aligned}
\langle \vec{x}, \vec{y} \rangle &= x_1 y_1 + \cdots + x_n y_n \\
&= y_1 x_1 + \cdots + y_n x_n \\
&= \langle \vec{y}, \vec{x} \rangle
\end{aligned} \quad (8)
$$

Scalar multiplication: If we scale either vector by a real number c, then the inner product will be scaled by c as well.

[1] In later classes, you'll see that this ceases to be true if we are working with complex numbers. But since we only work with real-valued vectors in this class, commutativity is always true for our purposes.

Proof:

$$
\begin{aligned}
\langle c\vec{x},\vec{y}\rangle &= (cx_1)y_1+\cdots+(cx_n)y_n \\
&= c(x_1y_1)+\cdots+c(x_ny_n) \\
&= c\,\langle \vec{x},\vec{y}\rangle
\end{aligned} \tag{9}
$$

By a similar argument (exercise: prove it yourself),

$$
\langle \vec{x},c\vec{y}\rangle = c\,\langle \vec{x},\vec{y}\rangle \tag{10}
$$

Distributive over vector addition: What happens when we take the inner product between a sum of two vectors and another vector? We can write:

$$
\begin{aligned}
\langle \vec{x}+\vec{y},\vec{z}\rangle &= (x_1+y_1)z_1+\cdots+(x_n+y_n)z_n \\
&= x_1z_1+y_1z_1+\cdots+x_nz_n+y_nz_n \\
&= x_1z_1+\cdots+x_nz_n+y_1z_1+\cdots+y_nz_n \\
&= \langle \vec{x},\vec{z}\rangle+\langle \vec{y},\vec{z}\rangle
\end{aligned} \tag{11}
$$

By a similar argument:

$$
\langle \vec{x},\vec{y}+\vec{z}\rangle = \langle \vec{x},\vec{y}\rangle+\langle \vec{x},\vec{z}\rangle \tag{12}
$$

21.5 Orthogonal Vectors

Two vectors $\vec{x}$, $\vec{y}$ in $\mathbb{R}^n$ are **orthogonal** if their inner product is zero, i.e. $\langle \vec{x},\vec{y}\rangle = 0$. The words perpendicular and orthogonal mean the same thing, but when working in dimensions greater than 2D or 3D, the term orthogonal tends to be more commonly used.

Here's an example, in $\mathbb{R}^3$. Let's say $\vec{x} = \begin{bmatrix}1\\1\\3\end{bmatrix}$ and $\vec{y} = \begin{bmatrix}4\\-1\\-1\end{bmatrix}$. Then, $\langle \vec{x},\vec{y}\rangle = (1)(4)+(1)(-1)+(3)(1) = 0$. Thus $\vec{x}$ and $\vec{y}$ are orthogonal.

Note that the standard unit vectors are always orthogonal to each other.

21.6 Special Vector Operations

The inner product is a basic building block for many operations. Here are some useful operations you can perform with the inner product calculation using different vectors as inputs.

These things are often important in computer programming contexts because computers (and programming languages) are often optimized to be able to do vector operations like inner products very fast. So, seeing an inner-product way of representing something can often speed up calculations considerably.

Sum of Components

$$\left\langle \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}^T, \begin{bmatrix} x_0 & x_1 & \cdots & x_n \end{bmatrix}^T \right\rangle = x_0 + x_1 + \ldots + x_n \tag{13}$$

Average

$$\left\langle \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \ldots & \frac{1}{n} \end{bmatrix}^T, \begin{bmatrix} x_0 & x_1 & \cdots & x_n \end{bmatrix}^T \right\rangle = \frac{x_0 + x_1 + \ldots + x_n}{n} \tag{14}$$

Sum of Squares

$$\left\langle \begin{bmatrix} x_0 & x_1 & \cdots & x_n \end{bmatrix}^T, \begin{bmatrix} x_0 & x_1 & \cdots & x_n \end{bmatrix}^T \right\rangle = x_0^2 + x_1^2 + \ldots + x_n^2 \tag{15}$$

Selective Sum Here, the first vector has values of 0's and 1's, where 1's correspond to the values of the second vector that are to be used in the sum. This, for example, can take the place of a for loop checking every element of a vector.

$$\left\langle \begin{bmatrix} 0 & 0 & 1 & 0 & 1 & \cdots & 0 & 1 \end{bmatrix}^T, \begin{bmatrix} x_0 & x_1 & \cdots & x_n \end{bmatrix}^T \right\rangle = x_2 + x_4 + \ldots + x_n \tag{16}$$

21.7 Norms

The **Euclidean Norm** of a vector is defined as:

$$\|\vec{x}\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} = \sqrt{\langle \vec{x}, \vec{x} \rangle} \tag{17}$$

Why is the norm important? The norm of a vector is also the magnitude of the vector (or length of the arrow). This corresponds to the usual notion of distance in $\mathbb{R}^2$ or $\mathbb{R}^3$. It is interesting to note that the set of points with equal Euclidean norm is a circle in $\mathbb{R}^2$ or a sphere in $\mathbb{R}^3$.

21.8 Properties of Norms

The following properties are true for all norms (not just the 2-norm).

Non-negativity: For $\vec{x}$ in $\mathbb{R}^n$,

$$\|\vec{x}\| \geq 0$$

Zero Vector: The zero vector is the only vector with a norm of zero. For $\vec{x}$ in $\mathbb{R}^n$,

$$\|\vec{x}\| = 0 \text{ if and only if } \vec{x} = \vec{0}$$

Scalar Multiplication: For vector $\vec{x}$ in $\mathbb{R}^n$ and scalar α,

$$\|\alpha\vec{x}\| = |\alpha|\|\vec{x}\|$$

Triangle Inequality: For vectors $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^n$,

$$\|\vec{x}+\vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$$

21.9 Interpretation of the Inner Product

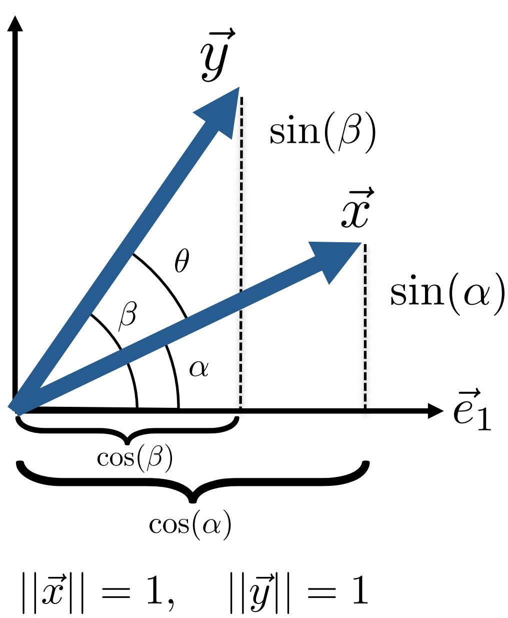
Now that we have defined the inner product, it's time to get an intuition about what an inner product is.

First we take the unit vector $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and a general unit vector in $\mathbb{R}^2$, $\vec{x} = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}$. $\vec{x}$ is a unit vector because $\|\vec{x}\| = \sqrt{\cos^2\alpha + \sin^2\alpha} = 1$.

When we draw the vectors, the angle between them is α. We can calculate

$$\begin{aligned}
\langle \vec{e}_1, \vec{x} \rangle &= \left\langle \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix} \right\rangle \\
&= (1)(\cos\alpha) + (0)(\sin\alpha) \\
&= \cos\alpha
\end{aligned}$$

This shows that the inner product of two unit vectors is the cosine of the angle between them when one of the vectors is pointed along the x-axis. Let's try to prove this for two general unit vectors in $\mathbb{R}^2$, $\vec{x} = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} \cos\beta \\ \sin\beta \end{bmatrix}$. (All unit vectors can be represented in this form).



Let's plug these vectors into the definition of the inner product. We'll use the trigonometric identity

$\cos\alpha\cos\beta + \sin\alpha\sin\beta = \cos(\alpha-\beta)$.

$$\begin{aligned}
\langle \vec{x}, \vec{y} \rangle &= \left\langle \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}, \begin{bmatrix} \cos\beta \\ \sin\beta \end{bmatrix} \right\rangle \\
&= \cos\alpha\cos\beta + \sin\alpha\sin\beta \\
&= \cos(\alpha-\beta) \\
&= \cos\theta
\end{aligned} \tag{18}$$

We've shown now that the inner product is the cosine of the angle between any two unit vectors.

What about any two general (nonzero) vectors $\vec{x}$ and $\vec{y}$? (Now we will let $\vec{x}$ and $\vec{y}$ be any length). We can first convert them into the unit vectors by dividing by the norm. This is called *normalization*.

$$\hat{x} = \frac{\vec{x}}{\|x\|} \qquad \hat{y} = \frac{\vec{y}}{\|y\|}$$

Since the direction of a vector does not change when it's multiplied by a scalar (and the norm is a scalar), then we know that the normalized vectors will point in the same direction as the original vectors. The angle between $\vec{x}$ and $\vec{y}$ is only dependent on their directions, so θ is unchanged by dividing by the norm. Therefore, we can apply Eq. (**??**) on the normalized vectors:

$$\left\langle \frac{\vec{x}}{\|x\|}, \frac{\vec{y}}{\|y\|} \right\rangle = \cos\theta \tag{19}$$

We can then use scalar multiplication property of inner products (Eq. (9)) since norms are scalars:

$$\begin{aligned}
\frac{\langle \vec{x}, \vec{y} \rangle}{\|x\|\|y\|} &= \cos\theta \\
\langle \vec{x}, \vec{y} \rangle &= \|x\|\|y\|\cos\theta
\end{aligned} \tag{20}$$

Now we have a geometric interpretation of the inner product: the inner product of vectors $\vec{x}$ and $\vec{y}$ is their lengths multiplied by the angle between them. One remarkable observation is that the inner product *does not depend on the coordinate system the vectors are in*, it only depends on the relative angle between these vectors and their length. This is the reason it is very useful in physics, where the physical laws do not depend on the coordinate system used to measure them. This is also the reason why this property holds in higher dimensions as well. For any two vectors we can look at the plane passing through them and the angle between them is the angle θ measured in the plane.

21.10 The Cauchy-Schwarz Inequality

The Cauchy-Schwarz inequality relates the inner product of two vectors to their length. The Cauchy-Schwarz inequality states:

$$\left| \langle \vec{x}, \vec{y} \rangle \right| \leq \|\vec{x}\|\|\vec{y}\| \tag{21}$$

We can prove this by recognizing that $|\cos\theta| \leq 1$ for all real θ. Thus:

$$\begin{aligned} |\langle\vec{x},\vec{y}\rangle| &= \Big| \|\vec{x}\|\|\vec{y}\|\cos\theta \Big| \\ &= \|\vec{x}\|\|\vec{y}\| \cdot |\cos\theta| \\ &\leq \|\vec{x}\|\|\vec{y}\| \end{aligned} \tag{22}$$

21.11 p-Norms [optional]

Sometimes, we write the Euclidean norm of a vector as $\|\vec{x}\|_2$, with a subscript 2. The subscript differentiates the Euclidean norm (or 2-norm) from other useful norms. In general, the p-norm is defined as:

$$\|\vec{x}\|_p = \left(|x_1|^p + |x_2|^p + \ldots + |x_n|^p\right)^{1/p} \tag{23}$$

These other norms might feel esoteric but turn out to be useful in many engineering settings. Follow-on courses like EE 127 and EE 123 will show you how these can be useful in applications like machine learning.

If no subscript is used, you can assume the 2-norm.

21.12 Practice Problems

These practice problems are also available in an interactive form on the course website.

1. We can verify that 2 vectors are orthogonal to each other by:
   (a) Multiplying their magnitudes together and checking if it is 1.
   (b) Taking the outer product of the two and checking if it is 0.
   (c) Multiplying the magnitudes and the cosine of the angle between them and checking if it is 0.
   (d) Multiplying the magnitudes and dividing by the cosine of the angle between them and checking if it is 1.
2. True or False: If two vectors $\vec{u}$ and $\vec{v}$ are orthogonal, then $\|\vec{u}+\vec{v}\| = \|\vec{u}\| + \|\vec{v}\|$.
3. True or False: If two vectors $\vec{u}$ and $\vec{v}$ are linearly dependent, then $|\langle\vec{u},\vec{v}\rangle| = \|\vec{u}\|\|\vec{v}\|$.
4. Does the following inner product definition (for $\mathbb{R}^2$) satisfy the all of the inner product properties? $\langle\vec{x},\vec{y}\rangle = \vec{x}^T \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \vec{y}$
5. Let $\vec{x},\vec{y} \in \mathbb{R}^n$ such that $\|\vec{x}\| = 1$. Find $\text{proj}_{\vec{x}}\vec{y}$.
   (a) $\frac{\langle\vec{y},\vec{x}\rangle}{\langle\vec{y},\vec{y}\rangle}\vec{x}$
   (b) $\frac{\langle\vec{y},\vec{x}\rangle}{\langle\vec{x},\vec{x}\rangle}\vec{y}$

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

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

6. What is the projection of the vector $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ onto the vector $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$?

7. Assume that you receive a shifted signal $\vec{r}$ and you correlate it with the original periodic signal $\vec{s}$, such that $\mathbf{C}_{\vec{s}}\vec{r} = \text{circcorr}(\vec{r}, \vec{s}) = \begin{bmatrix} -32 \\ -32 \\ 16 \\ 16 \\ 96 \\ 24 \end{bmatrix}$. What would you estimate as the sample delay between you and the transmitter?

8. True or False: $\|\mathbf{C}_{\vec{x}}\vec{y}\| = \left\|\mathbf{C}_{\vec{y}}\vec{x}\right\|$.

9. True or False: The projection of $\vec{y}$ onto $\vec{v}$ is the same as the projection of $\vec{y}$ onto $c\vec{v}$ whenever $c \neq 0$.