EECS 16B Fall 2019 — Designing Information Devices and Systems II — Discussion Worksheet

Discussion 7B

1 Gram-Schmidt Process

Gram-Schmidt is an algorithm that takes a set of linearly independent vectors $\{\vec{s}_1, \ldots, \vec{s}_n\}$ and generates an orthonormal set of vectors $\{\vec{q}_1, \ldots, \vec{q}_n\}$ that span the same vector space as the original set. Concretely, $\{\vec{q}_1, \ldots, \vec{q}_n\}$ satisfy the following:

- $\forall 0 < k \leq n$, span($\{\vec{s}_1, \ldots, \vec{s}_k\}$) = span($\{\vec{q}_1, \ldots, \vec{q}_k\}$)
- $\{\vec{q}_1, \ldots, \vec{q}_n\}$ is an orthonormal set of vectors

Definition: *Orthonormal*

A set of vectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is **orthonormal** if all the vectors are mutually orthogonal to each other and all are of unit length. That is:

- Orthogonal: For all pairs of vectors $\vec{v}_i, \vec{v}_j$ where $i \neq j$, $\langle \vec{v}_i, \vec{v}_j \rangle = 0$. For real vectors, this means $\vec{v}_i^T \vec{v}_j = 0$.
- Normalized: For all i, $\|\vec{v}_i\| = 1$. (This implies that $\|\vec{v}_i\| = \langle \vec{v}_i, \vec{v}_i \rangle = 1$.)

The Gram-Schmidt algorithm works by first finding the unit vector $\vec{q}_1$ such that span($\{\vec{q}_1\}$) = span($\{\vec{s}_1\}$). Subsequently, the unit vector $\vec{q}_2$ is calculated such that $\langle \vec{q}_1, \vec{q}_2 \rangle = 0$ and span($\{\vec{q}_1, \vec{q}_2\}$) = span($\{\vec{s}_1, \vec{s}_2\}$). This is continued through n vectors, resulting in the orthonormal set of vectors $\{\vec{q}_1, \ldots, \vec{q}_n\}$ that span the same vector space as $\{\vec{s}_1, \ldots, \vec{s}_n\}$.

How is this done? Finding $\vec{q}_1$ is straightforward, since it is the first vector in our new set, and therefore we must only satisfy $\|\vec{q}_1\| = 1$ and span($\{\vec{q}_1\}$) = span($\{\vec{s}_1\}$). Since span($\{\vec{s}_1\}$) is a one dimensional vector space, the unit vector that spans the same vector space would just be the unit vector in the same direction as $\vec{s}_1$. We have

$$\vec{q}_1 = \frac{\vec{s}_1}{\|\vec{s}_1\|}. \tag{1}$$

Calculating $\vec{q}_2$ requires that we satisfy:

(a) Spanning the same vector space as original set: span($\{\vec{q}_1, \vec{q}_2\}$) = span($\{\vec{s}_1, \vec{s}_2\}$)

(b) Orthogonal to previous vectors: $\langle \vec{q}_1, \vec{q}_2 \rangle = 0$

(c) Normalized: $\|\vec{q}_2\| = 1$

Using the vector $\vec{q}_1$ that we calculated above, we notice that

$$\text{span}(\{\vec{q}_1, \vec{s}_2\}) = \text{span}(\{\vec{s}_1, \vec{s}_2\}),$$

satisfying the first condition. However, $\vec{q}_1$ and $\vec{s}_2$ are not necessarily orthogonal.

We know from EE 16A that the following subspaces are equivalent for any pair of linearly independent vectors $\vec{v}_1, \vec{v}_2$:

- span($\vec{v}_1$, $\vec{v}_2$)
- span($\vec{v}_1$, $\alpha\vec{v}_2$)
- span($\vec{v}_1$, $\vec{v}_1 + \vec{v}_2$)
- span($\vec{v}_1$, $\vec{v}_1 - \vec{v}_2$)
- span($\vec{v}_1$, $\vec{v}_2 - \alpha\vec{v}_1$)

Let us choose vector

$$\vec{z}_2 = \vec{s}_2 - \alpha\vec{q}_1,$$

which will also have the same span as $\{\vec{q}_1, \vec{s}_2\}$ (and therefore the same span as $\{\vec{s}_1, \vec{s}_2\}$).

What should α be if we would like $\vec{q}_1$ and $\vec{z}_2 = \vec{s}_2 - \alpha\vec{q}_1$ to be orthogonal to each other? We know from working with Orthogonal Matching Pursuit (OMP), that $\vec{s}_2 - \text{proj}_{\vec{q}_1}(\vec{s}_2)$ will be orthogonal to $\vec{q}_1$, where

$$mboxproj_{\vec{q}_1}(\vec{s}_2) = \frac{\vec{q}_1^T\vec{s}_2}{\|\vec{q}_1\|^2}\vec{q}_1 = \vec{q}_1^T\vec{s}_2\vec{q}_1$$

is the projection of $\vec{s}_2$ onto $\vec{q}_1$. This makes sense because the projection of $\vec{s}_2$ onto $\vec{q}_1$ provides the component of $\vec{s}_2$ that is along $\vec{q}_1$. Subtracting off this component from $\vec{s}_2$ will only leave components of $\vec{s}_2$ that are orthogonal to $\vec{q}_1$.

Therefore, if we set

$$\alpha\vec{q}_1 = \text{proj}_{\vec{q}_1}(\vec{s}_2),$$

the resulting

$$\vec{z}_2 = \vec{s}_2 - \text{proj}_{\vec{q}_1}(\vec{s}_2) = \vec{s}_2 - \frac{\vec{q}_1^T\vec{s}_2}{\|\vec{q}_1\|^2}\vec{q}_1 = \vec{s}_2 - \vec{q}_1^T\vec{s}_2\vec{q}_1$$

will be orthogonal to $\vec{q}_1$.

To back up this intuition, let's solve for $\vec{z}_2$ algebraically using the definition of orthogonality:

$$\vec{q}_1^T\vec{z}_2 = 0 \tag{2}$$
$$\vec{q}_1^T(\vec{s}_2 - \alpha\vec{q}_1) = 0 \tag{3}$$
$$\vec{q}_1^T\vec{s}_2 - \alpha\|\vec{q}_1\|^2 = 0 \tag{4}$$
$$\alpha = \frac{\vec{q}_1^T\vec{s}_2}{\|\vec{q}_1\|^2} = \vec{q}_1^T\vec{s}_2 \tag{5}$$
$$\rightarrow \vec{z}_2 = \vec{s}_2 - \vec{q}_1^T\vec{s}_2\vec{q}_1 \tag{6}$$

Now we normalize $\vec{z}_2$ to complete the process of finding the $\vec{q}_2$ which satisfies all three conditions above:

$$\vec{q}_2 = \frac{\vec{z}_2}{\|\vec{z}_2\|}$$

In the question below, you will work through how this methodology leads to the Gram-Schmidt algorithm for calculating the orthonormal set $\{\vec{q}_1, \dots, \vec{q}_n\}$ from n linearly independent vectors $\{\vec{s}_1, \dots, \vec{s}_n\}$.

2 Stability

For some additional background on stability, see pages 9 through 15 of Prof. Murat Arcak's 16B Reader.

Suppose we have a discrete-time linear system with some disturbance, that is

$$\vec{x}(i+1) = A\vec{x}(i) + B\vec{u}(i) + \vec{w}(i) \tag{7}$$

in the general vector case. Until now, we have been essentially ignoring the disturbance $\vec{w}(i)$ out of convenience, that is treating the system (7) as if $\vec{w}(i) = 0$ for all i. However, we know that this is never true in practice! Wether it comes from true randomness in the system or through model imperfections, our system will always be affected by some disturbance.

What can disturbances do to our system? Well, the worst-case scenario is that the disturbance term $w(i)$ affects the system dynamics in such a way that the states $x(i)$ grow unboundedly with time. In most situations, having unbounded growth of the state is bad. Think of a car parked on top of a hill in neutral gear: a small push will cause its velocity to grow unboundedly, and the end result is decidedly not good.

This is where our notion of *stability* comes in. We will say that *a system is stable if $x(i)$ remains bounded for any initial condition and any bounded sequence of inputs $u(i)$*. This definition captures the intuitive notion that a small (i.e. bounded) disturbance should not be able to cause unbounded growth of the state.

2.1 Determining stability in discrete time systems

How can we tell if a system is stable or not? It would be nice it we could determine if a system is stable just from its representation (7). To start off with, we'll look at the scalar case, and see if we can figure out a rule. Consider the following three scalar systems:

- $x_1(i+1) = u_1(i) + w(i)$
- $x_2(i+1) = \frac{1}{2}x_2(i) + u_2(i) + w(i)$
- $x_3(i+1) = 2x_3(i) + u_3(i) + w(i)$

that are affected by some noise $w(i)$ that is *bounded*, that is $|w(i)| < \varepsilon$ for all times i, for some ε.

Suppose, furthermore, that all three have initial state $x_1(0) = x_2(0) = x_3(0) = x_0$, and based on this initial conditions we have designed input sequences $u_k(i)$ that bring the state to the origin in one time step. From $i = 1$ onwards, the state is only affected by the bounded disturbance. The question is, under these conditions, **which of the three states will remain bounded over time?**

The stability of these three systems, and your knowledge of sums of geometric series, should convince you that the following rule is valid: *for the scalar case $x(i+1) = \lambda x(i) + bu(i) + w(i)$, the system will be stable if $|\lambda| < 1$.*

For the vector case, a similar rule holds. Let λ be any particular eigenvalue of the matrix A in 7. Then *the vector system $\vec{x}(i+1) = A\vec{x}(i) + B\vec{u}(i) + \vec{w}(i)$ is stable is $|\lambda| < 1$ for all λ*. We will see why this is true later in the class.

1. **Gram-Schmidt Algorithm**

   Let's apply Gram-Schmidt orthonormalization to a set of three linearly independent vectors $\{\vec{s}_1, \vec{s}_2, \vec{s}_3\}$.

(a) Find unit vector $\vec{q}_1$ such that span($\{\vec{q}_1\}$) = span($\{\vec{s}_1\}$).

(b) Given $\vec{q}_1$ from the previous step, find $\vec{q}_2$ such that $span(\{\vec{q}_1, \vec{q}_2\}) = span(\{\vec{s}_1, \vec{s}_2\})$ and $\vec{q}_2$ is orthogonal to $\vec{q}_1$.

(c) Now given $\vec{q}_1$ and $\vec{q}_2$ in the previous steps, find $\vec{q}_3$ such that $span(\{\vec{q}_1, \vec{q}_2, \vec{q}_3\}) = span(\{\vec{s}_1, \vec{s}_2, \vec{s}_3\})$.

(d) Let's extend this algorithm to n linearly independent vectors. That is, given an input $\{\vec{s}_1, \ldots, \vec{s}_n\}$, write the algorithm to calculate the orthonormal set of vectors $\{\vec{q}_1, \ldots, \vec{q}_n\}$, where span($\{\vec{s}_1, \ldots, \vec{s}_n\}$) = span($\{\vec{q}_1, \ldots, \vec{q}_n\}$). *Hint: How would you calculate the i^{th} vector, $\vec{q}_i$?*

2. The Order of Gram-Schmidt

(a) If we are performing the Gram-Schmidt method on a set of vectors, does the order in which we take the vectors matter? Consider the set of vectors

$$\left\{ \vec{v}_1,\ \vec{v}_2,\ \vec{v}_3 \right\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \tag{8}$$

Perform Gram-Schmidt on these vectors first in the order $\vec{v}_1,\ \vec{v}_2,\ \vec{v}_3$. and then in the order $\vec{v}_3,\ \vec{v}_2,\ \vec{v}_1$. Do you get the same answer?

(b) Now perform Gram-Schmidt on these vectors in the order $\vec{v}_3,\ \vec{v}_2,\ \vec{v}_1$. Do you get the same result?

3. Orthonormal Matrices and Projections

An orthonormal matrix, $\mathbf{A}$, is a matrix whose columns, $\vec{a}_i$, are:

- Orthogonal (ie. $\langle \vec{a}_i, \vec{a}_j \rangle = 0$ when $i \neq j$)
- Normalized (ie. vectors with length equal to 1, $\|\vec{a}_i\| = 1$). This implies that $\|\vec{a}_i\|_2 = \langle \vec{a}_i, \vec{a}_i \rangle = 1$.

(a) Suppose that the matrix $\mathbf{A} \in \mathbb{R}^{N \times M}$ has linearly independent columns. The vector $\vec{y}$ in $\mathbb{R}^N$ is not in the subspace spanned by the columns of $\mathbf{A}$. What is the projection of $\vec{y}$ onto the subspace spanned by the columns of $\mathbf{A}$?

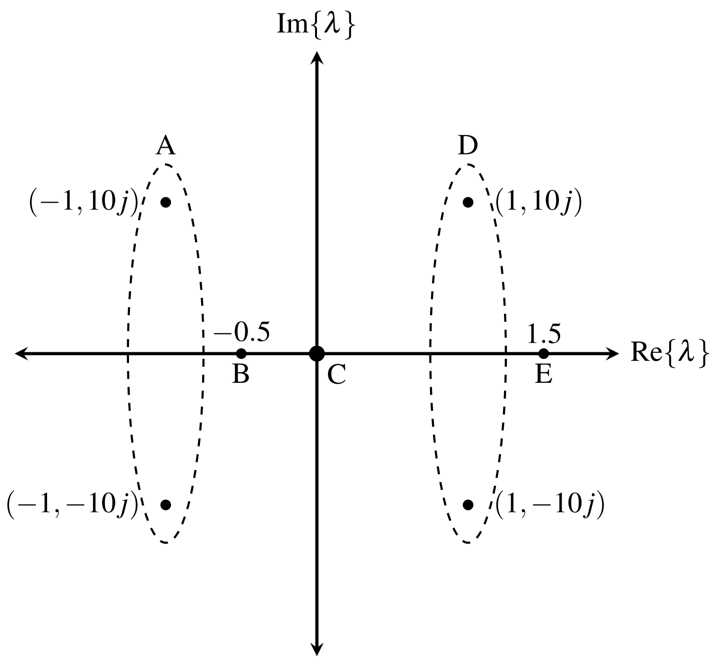
(b) Show if $\mathbf{A} \in \mathbb{R}^{N \times N}$ is an orthonormal matrix then the columns, $\vec{a}_i$, form a basis for $\mathbb{R}^N$.

(c) When $\mathbf{A} \in \mathbb{R}^{N \times M}$ and $N \geq M$ (i.e. tall matrices), show that if the matrix is orthonormal, then $\mathbf{A}^T\mathbf{A} = \mathbf{I}_{M \times M}$.

(d) Again, suppose $\mathbf{A} \in \mathbb{R}^{N \times M}$ where $N \geq M$ is an orthonormal matrix. Show that the projection of $\vec{y}$ onto the subspace spanned by the columns of $\mathbf{A}$ is now $\mathbf{A}\mathbf{A}^T\vec{y}$.

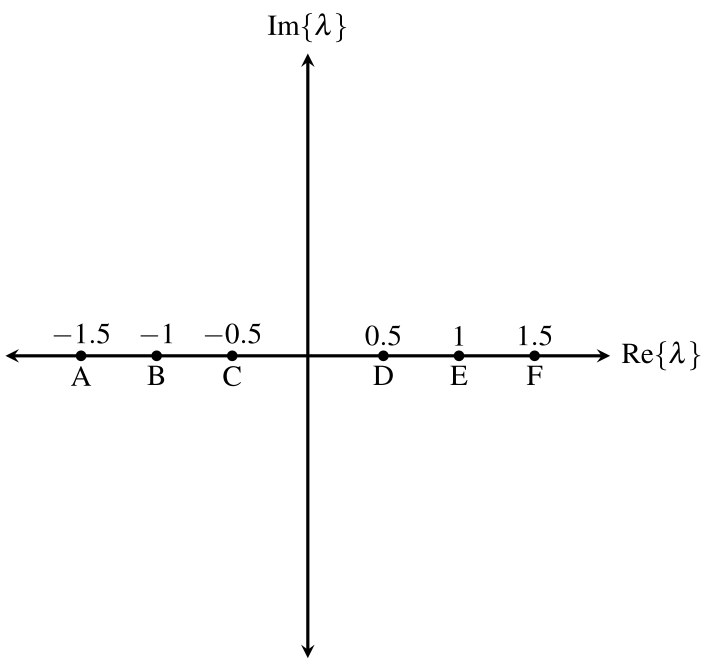
4. Continuous-time system responses

We have a system $\frac{d\vec{x}}{dt} = A\vec{x}$ with eigenvalues λ. For each set of λ values plotted on the real-imaginary axis, sketch $\vec{x}(t)$ with an initial condition of $x(0) = 1$. Determine if each system is stable.



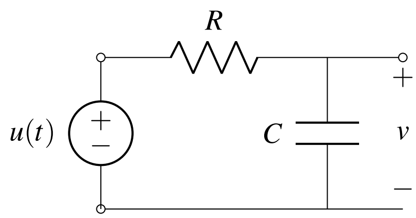
5. **Discrete time system responses**

We have a system $x(k+1) = \lambda x(k)$. For each λ value plotted on the real-imaginary axis, sketch $x(k)$ with an initial condition of $x(0) = 1$. Determine if each system is stable.



6. **Stability Examples and Counterexamples**

(a) Consider the circuit below with $R = 1\Omega$, $C = 0.5$F, and $u(t) = \cos(t)$. Furthermore assume that $v(0) = 0$ (that the capacitor is initially discharged).



This circuit can be modeled by the differential equation

$$\frac{d}{dt}v(t) = -2v(t) + 2u(t) \tag{9}$$

Show that the differential equation is always stable. Consider what this means in the physical circuit.

(b) Consider the system

$$x(t+1) = 2x(t) + u(t) \tag{10}$$

with $x(0) = 0$

Is the system stable or unstable? If unstable, find a bounded input sequence $u(t)$ that causes the system to 'blow up'. If unstable, is there still a (non-trivial) bounded input sequence that does not cause the system to 'blow up'?

Contributors:

- EE 16A folks.
- Regina Eckert.
- Sidney Buchbinder.
- Siddharth Iyer.
- Alex Devonport.
- Anant Sahai.

