1. **Controlling a System**

We are given a discrete time state space system, where $\vec{x}$ is our **state vector**, A is the **state space model** of the system, B is the **input matrix**, and $\vec{u}$ is the **control input vector**.

$$\vec{x}[t+1] = A\vec{x}[t] + B\vec{u}[t] \tag{1}$$

We define a system to be controllable if: **given a set of inputs, we can move the system from any initial state to any final state by choosing a sequence of inputs.** This has an important physical meaning; if a physical system is controllable, that means that we can get anywhere in the state space. For example, if a robot is controllable, it is able to travel anywhere in the system it is living in (given enough time and control inputs).

We will start with the assumption that the system started at rest at the zero vector: $\vec{x}[0] = \vec{0}$, and our state space will be an n-dimensional vector in $\mathbb{R}^n$.

(a) **How can you write out $\vec{x}[1]$ using the state space equation? How about $\vec{x}[2]$?**

(b) **Given these two observations and the initial condition $\vec{x}[0] = \vec{0}$, where can $\vec{x}$ reach in our state space reach after 2 time steps (i.e. at $t = 2$)?**

(c) **Given n observations, where in our state space can $\vec{x}$ reach at time $t = n$?**

(d) **Show that if the columns of A^kB are linearly dependent, then the columns of $A^{k+1}B$ are also linearly dependent.**

(e) To summarize our work from the previous parts, we now define a controllability matrix

$$\mathscr{C} = \begin{bmatrix} B & AB & A^2B & \cdots & A^{n-2}B & A^{n-1}B \end{bmatrix} \tag{2}$$

Based on the previous parts, when can we say our system is controllable? (i.e. we can reach anywhere in our state space using our system)

2. **Gram-Schmidt Process**

A set of vectors is orthonormal **if and only if**:

- Any pair of vectors are orthogonal (for any vectors $\vec{u}, \vec{v} \in S$ where $\vec{u} \neq \vec{v}$, the angle between them is $90°$).
- Each vector is of unit length (for vector $\vec{v} \in S$, $\|\vec{v}\| = 1$).

Gram-Schmidt orthonormalization is a process that allows us to take such a set S of arbitrary, linearly-independent vectors, and create a new set S' of vectors that span the exact same space, yet are now orthonormal.

(a) For linearly independent vectors $\vec{v}_1, \vec{v}_2$, **justify that Span($\vec{v}_1, \vec{v}_2$) is the same as Span($\vec{v}_1, \vec{v}_2 - \alpha\vec{v}_1$).**

(b) Let us iteratively generate the orthonormal set S' from S. Start by arbitrarily picking a vector $\vec{v}_1$ from S: **how can we create a new vector $\vec{u}_1$ for S' so that it spans the same space as the original vector, yet ensures that orthonormality is preserved for our new set?**

(c) Now, consider the next vector $\vec{v}_2$ in S. **How can we convert $\vec{v}_2$ to some $\vec{u}_2$ that is orthonormal with respect to the $\vec{u}_1$ that we've already added to S'?**

(d) **Write out the process for creating an arbitrary vector $\vec{u}_i$ for our orthonormal set S'.**

(e) Let $A = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 3 & 3 \\ -1 & -1 & 5 \\ 1 & 3 & 7 \end{bmatrix}$.

We will use Gram-Schmidt to find an orthonormal basis $U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{bmatrix}$ for the $\text{Col}(A)$.

i. **Find $\vec{u}_1$.**

ii. **Find $\vec{u}_2$.**

iii. **Find $\vec{u}_3$.**

3. Projection Matrices

We know from Gram-Schmidt, that if we have a basis $S = \{\vec{v}_1, \cdots \vec{v}_n\}$, for a vector space V, then we could pick an orthonormal basis $U = \{\vec{u}_1, \cdots \vec{u}_n\}$ for the same vector space, since we showed that the span of U was equivalent to the span of S, or the vector space V.

In this question, we will look into a special class of matrices called **projection matrices.** A projection matrix is a matrix P such that $P^2 = P$. Intuitively, if we projected a vector onto some subspace, then taking any vector in that subspace and reprojecting it will give the same vector.

(a) Let visit the least squares problem from 16A. Suppose we have some equation $A\vec{x} = \vec{b}$ that has no solution. Why does it have no solution? Because A has more rows than columns; it is a "tall" matrix where we have more equations than unknowns. How can we estimate the solution $\hat{x}$ using Least Squares?

(b) We can actually view the least squares problem as projection of the vector $\vec{b}$ onto the $\mathrm{Col}(A)$. What would the projection matrix P that projects $\vec{b}$ onto $\mathrm{Col}(A)$ be? Verify that it indeed is a projection matrix.

(c) From the result above, we see that we have to invert $A^T A$ every time. This takes a lot of computational power and is rather slow. If A were a matrix with orthonormal columns, how could this help us?

(d) We would like A to be orthonormal. We cannot, however, arbitrarily replace A with an orthonormal matrix Q. What must be true about Q for the least-squares equation to hold?

(e) We will create this orthonormal matrix Q by considering the columns of A as a basis for $\mathrm{Col}(A)$ and performing Gram-Schmdit on these basis vectors. We can then show that the projection of $\vec{b}$ onto $\mathrm{Col}(A)$ is the same as projecting it onto $\mathrm{Col}(Q)$ through the Orthogonal Decomposition Theorem. That is, $A(A^T A)^{-1} A^T \vec{b} = Q(Q^T Q)^{-1} Q^T \vec{b}$. What is the least square solution $\hat{x}$ in terms of Q?

