Discussion 11B

1. Minimum Energy Control & Spectral Theorem

In controllability/reachability analysis, we try to solve the linear system:

$$\mathcal{C}_{i^\star} \begin{bmatrix} \vec{u}[0] \\ \vdots \\ \vec{u}[i^\star - 1] \end{bmatrix} = \vec{x}^\star - A^{i^\star} \vec{x}_0 \tag{1}$$

for the vector quantities $\vec{u}[0], \ldots, \vec{u}[i^\star - 1]$. Cleaning up notation, let us fix $i^\star$, let $C := \mathcal{C}_{i^\star}$, let $\vec{z} := \vec{x}^\star - A^{i^\star} \vec{x}_0$, and let $\vec{u} := \begin{bmatrix} \vec{u}[0] \\ \vdots \\ \vec{u}[i^\star - 1] \end{bmatrix}$. Then this linear system becomes

$$C\vec{u} = \vec{z} \tag{2}$$

In the real world, we would like to use this framework to control mechanical systems, often expending the **minimum energy** possible.

But what do we mean by the "energy" of the control? In this context, we will use the squared norm of the input vector $\|u\|^2 = u_1^2 + \cdots + u_n^2$ as the model for the cost to applying a set of controls to our system.

Why do we use this definition? This formula for the cost model is closely connected to many physical scenarios relating to energy. Consider a few examples below:

- $E_{\text{capacitor}} = \frac{1}{2}CV^2$
- $E_{\text{spring}} = \frac{1}{2}kx^2$
- $E_{\text{kinetic}} = \frac{1}{2}mv^2$

And so we find that the definition we use is a natural one.

Optional EECS16A Refresher: Recall the following vector spaces:

The range (or column space) of a matrix A refers to the following vector space $\mathrm{Col}(A) = \{A\vec{x} : \vec{x} \in \mathbb{R}^n\}$*. It is the vector space consisting of all possible linear combinations of the columns of A.*

Then, there is the null space of A, which refers to the following vector space $\mathrm{Null}(A) = \{\vec{x} : A\vec{x} = 0\}$.

(a) Suppose we would like to solve for the minimum energy (or minimum norm) solution of the linear system $C\vec{u} = z$. This problem can be expressed as the following optimization problem:

$$\underset{\vec{u}}{\operatorname{argmin}} \|\vec{u}\|^2 = \underset{u[i]}{\operatorname{argmin}} \sum_{i=0}^{\ell-1} u[i]^2 \tag{3}$$

$$\text{s.t.} \quad C\vec{u} = \vec{z} \tag{4}$$

Symmetric matrices can make calculations and reasoning about the properties of a matrix much easier. Suppose C is a *real, symmetric matrix*. **Rewrite C in terms of its spectral decomposition** (take Q to be the orthonormal basis of eigenvectors of C and Λ to be the diagonal matrix of the eigenvalues).

Solution: We recall from Spectral Theorem that a symmetric matrix can be written as:

$$C = Q\Lambda Q^T \tag{5}$$

(b) One way to look at minimum energy control is through lens of our vector spaces. How might we do this?

Well, Spectral Theorem tells us that Q is an orthonormal basis of $\mathbb{R}^n$. If $\text{Rank}(C) = r$, then Q can be written as the block matrix $\begin{bmatrix} Q_r & Q_{n-r} \end{bmatrix}$ where Q_r forms an orthonormal basis for $\text{Col}(C)$ and Q_{n-r} similarly forms one for $\text{Null}(C)$.

Let's perform an orthonormal basis change:

$$\vec{u} = Q\tilde{\vec{u}} \tag{6}$$

Using our new basis, rewrite $\vec{u}$ in terms of Q_r and Q_{n-r}.

(HINT: Consider breaking up Q and $\vec{u}$ into a block matrix and partitioned vector respectively.) **Solution:** From the information given in the problem, we can write the orthonormal basis change as a block matrix multiplied to a partitioned vector as follows:

$$\vec{u} = Q\tilde{\vec{u}} \tag{7}$$

$$= \begin{bmatrix} Q_r & Q_{n-r} \end{bmatrix} \begin{bmatrix} \tilde{\vec{u}}_{\text{Col}(C)} \\ \tilde{\vec{u}}_{\text{Null}(C)} \end{bmatrix} \tag{8}$$

$$= Q_r \tilde{\vec{u}}_{\text{Col}(C)} + Q_{n-r} \tilde{\vec{u}}_{\text{Null}(C)} \tag{9}$$

Where $\tilde{\vec{u}}_{\text{Col}(C)}$ are the entries in $\tilde{\vec{u}}$ that correspond to the Q_r columns and $\tilde{\vec{u}}_{\text{Null}(C)}$ are those that correspond with Q_{n-r}

(c) Ultimately, the objective we are trying to minimize is still $\|\vec{u}\|^2$. **Use your findings from part (b) to show that $\|\vec{u}\|^2 = \|\tilde{\vec{u}}_{\text{Col}(C)}\|^2 + \|\tilde{\vec{u}}_{\text{Null}(C)}\|^2$.**

(HINT: Given some arbitrary orthonormal matrix U and arbitrary vector $\vec{p}$, how are $\|\vec{p}\|$ and $\|U\vec{p}\|$ related?)

Solution: We start by using what we found in part (b) to find $\|\vec{u}\|^2$:

$$\begin{aligned}
\|\vec{u}\|^2 &= \|Q_r\widetilde{\vec{u}}_{\text{Col}(C)} + Q_{n-r}\widetilde{\vec{u}}_{\text{Null}(C)}\|^2 && (10)\\
&= \langle Q_r\widetilde{\vec{u}}_{\text{Col}(C)} + Q_{n-r}\widetilde{\vec{u}}_{\text{Null}(C)}\rangle && (11)\\
&= (Q_r\widetilde{\vec{u}}_{\text{Col}(C)} + Q_{n-r}\widetilde{\vec{u}}_{\text{Null}(C)})^T(Q_r\widetilde{\vec{u}}_{\text{Col}(C)} + Q_{n-r}\widetilde{\vec{u}}_{\text{Null}(C)}) && (12)\\
&= (Q_r\widetilde{\vec{u}}_{\text{Col}(C)})^T(Q_r\widetilde{\vec{u}}_{\text{Col}(C)}) + 2(Q_r\widetilde{\vec{u}}_{\text{Col}(C)})^T(Q_{n-r}\widetilde{\vec{u}}_{\text{Null}(C)}) + (Q_{n-r}\widetilde{\vec{u}}_{\text{Null}(C)})^T(Q_{n-r}\widetilde{\vec{u}}_{\text{Null}(C)}) && (13)\\
&= \widetilde{\vec{u}}_{\text{Col}(C)}^T Q_r^T Q_r \widetilde{\vec{u}}_{\text{Col}(C)} + 2\widetilde{\vec{u}}_{\text{Col}(C)}^T Q_r^T Q_{n-r}\widetilde{\vec{u}}_{\text{Null}(C)} + \widetilde{\vec{u}}_{\text{Null}(C)}^T Q_{n-r}^T Q_{n-r}\widetilde{\vec{u}}_{\text{Null}(C)} && (14)\\
&= \widetilde{\vec{u}}_{\text{Col}(C)}^T\widetilde{\vec{u}}_{\text{Col}(C)} + \widetilde{\vec{u}}_{\text{Null}(C)}^T\widetilde{\vec{u}}_{\text{Null}(C)} && (15)\\
&= \|\widetilde{\vec{u}}_{\text{Col}(C)}\|^2 + \|\widetilde{\vec{u}}_{\text{Null}(C)}\|^2 && (16)
\end{aligned}$$

Where we use orthonormality to conclude that $Q_r^T Q_r = I_{r\times r}$, $Q_{n-r}^T Qn - r = I_{(n-r)\times(n-r)}$, and $Q_r^T Q_{n-r} = \vec{0}_{r\times(n-r)}$.
Thus, $\|\vec{u}\|^2 = \|\widetilde{\vec{u}}_{\text{Col}(C)}\|^2 + \|\widetilde{\vec{u}}_{\text{Null}(C)}\|^2$.

(d) Putting everything from the last two parts together, let us now solve for the solution to the minimum energy problem (written once more for convenience):

$$\underset{\vec{u}}{\text{argmin}}\,\|\vec{u}\|^2 = \underset{u[i]}{\text{argmin}} \sum_{i=0}^{\ell-1} u[i]^2 \quad (17)$$

$$\text{s.t.} \quad C\vec{u} = \vec{z} \quad (18)$$

Solve for the optimal minimum energy input $\vec{u}^*$ in its simplest form in terms of $\vec{u}_{\text{Col}(C)}$ and/or $\vec{u}_{\text{Null}(C)}$. Explain what your result means intuitively.
(HINT: Which of $\vec{u}_{\text{Col}(C)}$ or $\vec{u}_{\text{Null}(C)}$ doesn't effect $C\vec{u}$ (try to think about your vector space definitions)? What should we do to that value if we want to minimize the squared norm of the input?)

Solution: Remember that we can rewrite the squared norm of the input (the energy, which is our optimization objective) as follows:

$$\|\vec{u}\|^2 = \|\widetilde{\vec{u}}_{\text{Col}(C)}\|^2 + \|\widetilde{\vec{u}}_{\text{Null}(C)}\|^2 \quad (19)$$

Based on how we defined $\widetilde{\vec{u}}_{\text{Null}(C)}$, we know it has no effect on $C\vec{u}$. Thus, we should minimize its affect on the squared norm of the input by *setting it to zero*. Thus, our solution is as follows:

$$\min\|\vec{u}\|^2 = \min(\|\widetilde{\vec{u}}_{\text{Col}(C)}\|^2 + \|\widetilde{\vec{u}}_{\text{Null}(C)}\|^2) = \|\widetilde{\vec{u}}_{\text{Col}(C)}\|^2 \quad (20)$$

Where the optimial min. energy solution $\vec{u}^* = \widetilde{\vec{u}}_{\text{Col}(C)}$ such that $\widetilde{\vec{u}}_{\text{Null}(C)} = \vec{0}$. What does this mean intuitively? Essentially, the min. energy solution *ignores or zeros out columns of the controllability matrix that don't help us get closer to our desired state.*
Let's clarify a bit as to why $\widetilde{\vec{u}}_{\text{Null}(C)}$ doesn't affect $C\vec{u}$. We can rewrite $C\vec{u}$ as follows:

$$C\vec{u} = CQ_r\widetilde{\vec{u}}_{\text{Col}(C)} + CQ_{n-r}\widetilde{\vec{u}}_{\text{Null}(C)} \quad (21)$$

But recall that $Q_{n-r}\widetilde{\vec{u}}_{\text{Null}(C)}$ lives in $\text{Null}(C)$! Thus, $CQ_{n-r}\widetilde{\vec{u}}_{\text{Null}(C)} = \vec{0}$ and doesn't affect $C\vec{u}$.

(e) Now, let's do a numerical example. Consider the following linear discrete time system

$$\vec{x}[i+1] = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \vec{x}[i] + \begin{bmatrix} -1 \\ 1 \end{bmatrix} u[i] \tag{22}$$

Find the controllability matrix C for this system.

Solution: We solve for the controllability matrix exactly as shown in lecture in the past:

$$C = \begin{bmatrix} A\vec{b} & \vec{b} \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{23}$$

(f) Now, suppose we want to achieve desired state of $\vec{x}[2] = \begin{bmatrix} 4 \\ -4 \end{bmatrix}$ at timestep $i = 1$. Assume your initial condition is $\vec{x}[0] = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Write your linear system to solve for the input vector $\vec{u} = \begin{bmatrix} \vec{u}[0] \\ \vec{u}[1] \end{bmatrix}$ in the form $C\vec{u} = \vec{z}$. Then, solve for one $\vec{u}$ that achieves the desired system state. Remember, there will be many solutions as the system is underdetermined.

(HINT: Make use of the linear system formulation that comes as a result of controllability analysis shown on page 1.)

Solution: We have already found our controllability matrix C. All that is left is to find $\vec{z}$. From our controllability form, we know that $\vec{z} = \vec{x}^\star - A^{i^\star}\vec{x}_0$.

$$\vec{z} = \vec{x}^\star - A^{i^\star}\vec{x}_0 = \vec{x}[2] = \begin{bmatrix} 4 \\ -4 \end{bmatrix} \tag{24}$$

Therefore, our linear system to solve for $\vec{u}$ becomes:

$$\begin{bmatrix} A\vec{b} & \vec{b} \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \vec{u} = \begin{bmatrix} 4 \\ -4 \end{bmatrix} \tag{25}$$

Since this system is underdetermined, a solution for $\vec{u}$ will be such that $\vec{u}[0] - \vec{u}[1] = 4$. $\begin{bmatrix} 2 \\ -2 \end{bmatrix}$, $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$, and $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$ are all possible answers students could give.

(g) Finally, notice that $\mathrm{Col}(C) = \mathrm{span}\left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ and $\mathrm{Null}(C) = \mathrm{span}\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$.

The minimum norm solution is $\vec{u}_{\min} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$. We will compare this to another arbitrary solution $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$.

i **Write both $\vec{u}_{\text{min}}$ and $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ as a linear combination of the column and null space span vectors. Compare the coefficients of the null space span vector.** **Solution:** This can be done through simple inspection.

For $\vec{u}_{\text{min}}$:

$$\vec{u}_{\text{min}} = 2 \times \begin{bmatrix} 1 \\ -1 \end{bmatrix} + 0 \times \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{26}$$

For $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$:

$$\begin{bmatrix} 3 \\ -1 \end{bmatrix} = 2 \times \begin{bmatrix} 1 \\ -1 \end{bmatrix} + 1 \times \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{27}$$

We notice that the coefficient of the null space vector for the minimum energy/norm solution is zero, aligning with what we found earlier in the problem.

ii **Compare the norms of the two solutions. Verify that $\|\vec{u}_{\text{min}}\|$ is smaller.** **Solution:** We compare $\vec{u}_{\text{min}}$ to $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$:

$$\begin{aligned} \|\vec{u}_{\text{min}}\| &= \sqrt{2^2 + (-2)^2} && (28) \\ &= \sqrt{8} && (29) \\ & && (30) \end{aligned}$$

$$\begin{aligned} \left\| \begin{bmatrix} 3 \\ -1 \end{bmatrix} \right\| &= \sqrt{3^2 + (-1)^2} && (31) \\ &= \sqrt{10} && (32) \\ & && (33) \end{aligned}$$

Thus, $\|\vec{u}_{\text{min}}\| < \left\| \begin{bmatrix} 3 \\ -1 \end{bmatrix} \right\|$

Contributors:

- Anant Sahai.
- Nathan Lambert.
- Kareem Ahmad.
- Neelesh Ramachandran.
- Chancharik Mitra.
- Nikhil Jain.