EECS 16B Designing Information Devices and Systems II
Fall 2021 Note 11: Controllability

1 Overview

In note 9, system identification was described, wherein a continuous time system was modeled with a discrete time system by learning from data. In note 10, the last section of closed-loop control described a system we were able to stabilize by feedback control, i.e. $u[i] = K\vec{x}[i]$. However, the questions of "can a system always be stabilized?" and "can we plan to reach a state in a certain number of time steps and achieve it with an appropriate set of controls?" remain unanswered. We care to answer these questions in the setting of our systems that we have learned from data and characterized the stability of, especially when they are unstable and cannot reject disturbances. It turns out that these two questions can be answered by discussion of the notion of controllability of a system: whether our system can or cannot be driven to a specific state. The discussion of how to refine our plans and reach our goal state in an elegant way (e.g. in consideration of a notion of limited resources or at a certain rate) will be covered later.

2 Special Cases of Controllability

Many of the systems that we study have the following form:

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

We want to use our observations of $\vec{x}[t]$ and apply some *control* to the system (via the input $\vec{u}[t]$) to make $\vec{x}[t]$ approach some target $\vec{x}^\star$ over time. Under what conditions can we do this? In this model, our input vector can influence the next timestep's state vector directly. Not only that, but comparing the dimensions of the quantities in our equation, the input vector has as many elements as the state vector itself.

So, if we have full control over $\vec{u}[t]$, we are able to directly solve this problem! Given some $\vec{x}[0]$, we may choose $\vec{u}[0] = \vec{x}^\star - A\vec{x}[0]$, to drive $\vec{x}[1]$ to our desired state in a single time step. This is because we are setting the input to be the difference between where we start and where we want to end up.

However, in reality, we seldom have that many degrees of freedom in our input $\vec{u}[t]$. Imagine, for instance, that our state equation modeled the behavior of a circuit connected to an input voltage that we can vary. We have seen previously that the state vector of a circuit may include many components, like the voltages across capacitors or the currents through inductors, none of which we can directly change. Therefore, we use the following equation to obtain a more accurate model of our system:

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

This equation acknowledges that the number of inputs we can independently adjust at any single time might be different from the dimension of the state. Notice that we have added the matrix B, which constrains how our input $\vec{u}[t]$ can influence the evolution of our system over time. For instance, in a system with a two-dimensional state, the case of

$$B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

would mean that our input can only directly affect the second state. For now, we will treat the above discrete-time equation as an accurate model of our system (ignoring noise and other disturbances from Note 7B).

Before, we were always able to choose a $\vec{u}[0]$ to get to our target state in a single time step, no matter what the initial state $\vec{x}[0]$ was. Let's try doing the same thing now for an arbitrary B:

$$\begin{aligned} & \vec{x}[1] = \vec{x}^{\star} \\ \implies \quad & A\vec{x}[0] + B\vec{u}[0] = \vec{x}^{\star} \\ \implies \quad & B\vec{u}[0] = \vec{x}^{\star} - A\vec{x}[0]. \end{aligned}$$

Observe here that we can only achieve our desired state if $\vec{x}^{\star} - A\vec{x}[0] \in \text{range}(B)$, which is not always the case. For example, B may be a "tall" matrix (more rows than columns). Let n be the dimension of $\vec{x}$ - in other words, let $\vec{x}$ be made up of n scalar components. For *any* desired state to be reachable, B would have to span all of $\mathbb{R}^n$, and so be of rank n.

So, is all hope lost? Not necessarily. Recall that we only wanted to achieve our target state at *some point*, not necessarily in a single time step. Consider the state transition matrix

$$A = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix},$$

with B defined as before (i.e. only allowing us to affect the second state). Since B has only one column, our control input $\vec{u}$ is one dimensional, and so can be treated as a scalar u. We know that, after 1 time step, we can only reach the states that can be written as

$$\vec{x}[1] = A\vec{x}[0] + Bu[0]$$

for a suitable choice of control $u[0]$. Since B is not of rank 2, the set of viable $\vec{x}[1]$ does not yet span all of $\mathbb{R}^2$. However, now consider the set of states that are reachable after 2 time steps. By our state transition equation, they can be written as

$$\begin{aligned} \vec{x}[2] &= A\vec{x}[1] + Bu[1] \\ &= A\left(A\vec{x}[0] + Bu[0]\right) + Bu[1] \\ &= A^2\vec{x}[0] + ABu[0] + Bu[1] \end{aligned}$$

Notice that if the vectors AB and B together span all of $\mathbb{R}^2$, then we would be able to choose coefficients $u[0]$ and $u[1]$ (that is, two different controls at timesteps 0 and 1) to reach *any desired* $\vec{x}[2]$, meaning that the system is controllable in 2 time steps.

3 General Controllability

Generalizing our equation for $\vec{x}[2]$ to the case of multidimensional control inputs (i.e. when B has k columns, not just 1, and our input has multiple components), in addition to arbitrarily large state dimensions n, we obtain

$$\vec{x}[2] = A^2\vec{x}[0] + AB\vec{u}[0] + B\vec{u}[1]$$

from an analogous calculation.

This can be re-expressed as

$$\vec{x}[2] = A^2\vec{x}[0] + \begin{bmatrix} AB & B \end{bmatrix} \begin{bmatrix} \vec{u}[0] \\ \vec{u}[1] \end{bmatrix},$$

using stacked matrix notation.

More generally, we can show that as time passes, and the effect of our control inputs propagate, the state vector after n timesteps becomes:

$$\vec{x}[t] = A^n\vec{x}[0] + \begin{bmatrix} B & AB & \cdots & A^{t-1}B \end{bmatrix} \begin{bmatrix} \vec{u}[t-1] \\ \vdots \\ \vec{u}[0] \end{bmatrix}.$$

This pattern is familiar to us from our analysis of discretization, and the way we unrolled the recursion in an analogous way in dis02B and dis06A. We have chosen to order our inputs in "reverse" order, such that the most recently applied input comes on top. This structure allows us to present the stacked matrix in a more natural order (B is the effect of the most recent input, AB is the effect of the input before that, and so on). Therefore, we find that our system is controllable after t time steps if and only if the matrix

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

has a column span that is n-dimensional.

The question now becomes - is *every* system controllable after sufficiently many time steps? That is to say, is it true that, for any $n \times n$ matrix A and $n \times k$ matrix B, there exists some constant t such that the matrix $\begin{bmatrix} B & AB & \cdots & A^{t-1}B \end{bmatrix}$ is of full rank?

Unfortunately, this is not the case - we can construct a counterexample by considering the case when the input has no effect on the state transition, and $B = 0$! But even with nontrivial B matrices, some systems may not be controllable. Consider, for instance:

$$A = \begin{bmatrix} -2 & 2 \\ 2 & 1 \end{bmatrix}$$
$$B = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Observe that

$$AB = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$
$$A^2B = \begin{bmatrix} 4 \\ 8 \end{bmatrix},$$

and so on, so all of the columns of the stacked matrix will be linearly dependent. No matter how many timesteps we go, columns of this form will not add to our column span. Thus, even though this system has a nontrivial state transition equation, it is still not controllable even after infinitely many time steps. Note that the reason why we are observing a lack of increase in the span of our space in this example is because our B matrix is an eigenvector of our A matrix. Similarly, we could imagine that for a system with A matrix, $n \times n, n > 2$, and B matrix with $m < n$ eigenvector columns, or linear combinations of eigenvector columns, will similarly fail to be controllable.

4 Determining Controllability

We will now attempt to devise a general method of determining whether or not a system is controllable. In essence, we want to know whether the infinite sequence

$$\text{range}\left(\begin{bmatrix} B \end{bmatrix}\right), \text{range}\left(\begin{bmatrix} B & AB \end{bmatrix}\right), \text{range}\left(\begin{bmatrix} B & AB & A^2B \end{bmatrix}\right), \ldots$$

will ever reach $\mathbb{R}^n$. If, at some finite point, this sequence reaches $\mathbb{R}^n$, then that's great, and we know that our system is controllable! However, if our system turns out to be uncontrollable, then this sequence will keep going on forever and will never reach $\mathbb{R}^n$. Unfortunately, it is not possible to somehow evaluate an infinite number of terms of our sequence. So then, with only a finite number of terms of the sequence, how can we ever confidently say: "No, this system is not controllable"?

Our approach will rely on the following key observation: that if all the columns of A^tB are linearly dependent on the previous A^kB for $k < t$, then all the columns of $A^{t+1}B$ will also be linearly dependent on the same set of A^kB. That is to say, if the ranges in the above sequence *ever stop growing* for even a single iteration, then they will *never start growing again*. Making and verifying such observations of the inheritance of such properties on the iteration of some process or variable is the hallmark of the proof method induction, which you will learn about in more detail in CS70.

For simplicity, we will conduct our proof under the assumption that $B = \vec{b}$ is a column vector, so all the $A^tB = A^t\vec{b}$ have only one column. However, the exact same approach works in the most general case, except that the notation is more messy.

By the condition of the key observation, and the definition of linear dependence, there exist coefficients α_k such that:

$$A^t\vec{b} = \overbrace{\sum_{k=0}^{t-1} \alpha_k A^k \vec{b}} \tag{1}$$

linear combination of previous vectors

Thus, we may write the following sequence of equations:

$$\begin{aligned}
& & A^{t+1}\vec{b} &= A(A^t\vec{b}) \\
&\text{[by eq. (1)]} & &= A\sum_{k=0}^{t-1} \alpha_k A^k \vec{b} \\
&\text{[bring } A \text{ inside sum]} & &= \sum_{k=0}^{t-1} \alpha_k A^{k+1} \vec{b} \\
&\text{[re-number indices]} & &= \sum_{k=1}^{t} \alpha_{k-1} A^{k} \vec{b}
\end{aligned}$$

$$\begin{aligned}
&\text{[split sum; separate } k=t \text{ case]} && = \alpha_{t-1}A^t\vec{b} + \sum_{k=1}^{t-1}\alpha_{k-1}A^k\vec{b} \\
&\text{[sub. eq. (1), bring in scalar]} && = \sum_{k=0}^{t-1}\alpha_{t-1}\alpha_k A^k\vec{b} + \sum_{k=1}^{t-1}\alpha_{k-1}A^k\vec{b} \\
&\text{[group terms, separate } k=0\text{]} && = \alpha_{t-1}\alpha_0\vec{b} + \sum_{k=1}^{t-1}(\alpha_{t-1}\alpha_k + \alpha_{k-1})A^k\vec{b}
\end{aligned}$$

So, we have shown that $A^{t+1}B$ can be expressed as a linear combination of the same preceding A^kB.

With this result, we are on the road to developing an algorithm to determine the controllability of a system. We iteratively construct the sequence of ranges as described above. If the dimension of the ranges ever stops growing, we know that the dimension will never start to grow again as the sequence continues, so we know what possible $\vec{x}^\star$ are reachable even as $t \to \infty$. The only issue is if the dimension of the ranges never stops growing. But this can never be the case, since the dimension of the ranges is bounded by n (since they are all subspaces of $\mathbb{R}^n$)!

Thus, in summary: After n iterations, the ranges will either have stopped growing during one of the intermediate iterations, or would have reached rank n and so will not be able to continue to grow thereafter.

The argument above was given for any single column $\vec{b}$ and so in spirit, it can apply to all the columns viewed one at a time. So, we might as well do them all together. Consequently, by considering the subspace

$$\text{range}\left(\begin{bmatrix} B & AB & \dots & A^{n-1}B \end{bmatrix}\right)$$

we will be able to determine whether our system is controllable. If this range is of dimension n (and so spans the entirety of $\mathbb{R}^n$), then our system is controllable. Otherwise, it is not. Typically, we refer to this span as range($\mathscr{C}$), where $\mathscr{C}$ is defined as

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

An immediate consequence of this result is that a system is controllable *if and only if* it is controllable in at most n timesteps.

The next natural question is as follows: even if our system is controllable, how do we calculate the necessary inputs needed to reach a desired state $\vec{x}^\star$? As it turns out, this is straightforward, if we consider the stacked matrix representation of $\vec{x}[n] = \vec{x}^\star$:

$$\vec{x}^\star = A^n\vec{x}[0] + \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}\begin{bmatrix} \vec{u}[n-1] \\ \vdots \\ \vec{u}[0] \end{bmatrix} = A^n\vec{x}[0] + \mathscr{C}\begin{bmatrix} \vec{u}[n-1] \\ \vdots \\ \vec{u}[0] \end{bmatrix}$$

Rearranging, we obtain

$$\mathscr{C}\begin{bmatrix} \vec{u}[n-1] \\ \vdots \\ \vec{u}[0] \end{bmatrix} = \vec{x}^\star - A^n\vec{x}[0].$$

Using Gaussian elimination, we can now determine a sequence of control inputs to use, if such a sequence exists. Thus, we have resolved the problem that was initially posed at the beginning of our discussion. We have a condition on whether a system can reach a desired state. We have implicitly also answered the question of whether a system is stabilizable, as if we know our system is controllable, we can drive the system to the state $\vec{0}$ in a number of time steps, and the system state will remain at $\vec{0}$ with no further input, provided we have no disturbance. In the presence of disturbances, we can continue to do this repeatedly with bounded inputs, and we will remain within a region close to $\vec{0}$. However, we have said nothing of whether we can do so at specific rate i.e. if we can always set the eigenvalues of the system to make our system intrinsically disturbance rejecting. Later, we will also learn how to choose potentially better sequences of control inputs if the solution above is not unique.

As a final note, we have discussed controllability in the context of discrete time - the corresponding concept exists for continuous time systems, but we do not discuss it here. It turns out it is the same test - we check the rank of $\mathscr{C}$ for system matrices A and B in $\frac{\mathrm{d}}{\mathrm{d}t}\vec{x} = A\vec{x} + B\vec{u}$. A sketch for why this is true is because a continuous time system can discretized. If this discretized version is controllable, then the underlying continuous time system is also controllable.

Contributors:

- Moses Won.
- Neelesh Ramachandran.
- Rahul Arya.
- Anant Sahai.