EECS 16A   Designing Information Devices and Systems I
# Fall 2018   Discussion 6A

1. Justin Beaver (Fall 2015 MT1)

In your homework, there was a question about Justin Bieber's segway — that was about controlling a multi-dimensional system with one control input. In this problem, we will instead think about a curious and superintelligent beaver watching the water level in a pool — this is implicitly about how many sensors are needed to measure the state of a multi-dimensional system.

Three superintelligent rodents live in a network of pools. Justin Beaver lives in pool 1, Selena Gopher lives in pool 2, and Mousey Cyrus lives in pool 3. They are sadly not on talking terms, but Justin really wants to know about the other pools.

Suppose there is a network of pumps connecting the three different pools, given in the figure. $x_1[t], x_2[t]$, and $x_3[t]$ is the water level in each pool at time step t. At each time step, the water from each pool is pumped along the arrows. The water levels are updated according to the matrix:
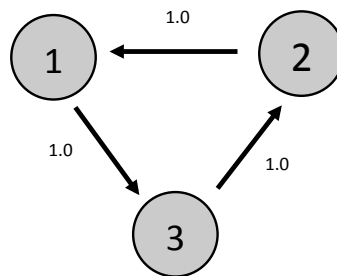
$$\begin{bmatrix} x_1[t+1] \\ x_2[t+1] \\ x_3[t+1] \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} x_1[t] \\ x_2[t] \\ x_3[t] \end{bmatrix}$$

(a) Justin lives in pool 1 so he can watch the water level in this pool. He also knows exactly how the pumps work — i.e. knows the pump matrix **A**. **Can Justin figure out the initial water levels in all three pools just by watching the water levels in his pool as time goes by? Describe (briefly) in words how to do this. How many times does Justin need to observe the water in his own pool to figure this out?**

(Hint: No "linear algebra" machinery is needed here. Just think about what Justin observes as time goes by.)

Answer:

Because the water flow from one pool into another is fairly simple in this system, we can figure out intuitively that Justin will be able to figure out the initial water level in each pool. (We will see later that in situations where the system of pumps is more complex, we will need to be much more careful. There are some more complex systems of pumps that will be able to effectively hide information about water levels in the other pools from Justin.)

In this case, however, Justin's analysis is straightforward. At each time step, all the water from pool 1 flows into pool 3, all the water from pool 2 flows into pool 1, and all the water from pool 3 flows into pool 2. Because of the cyclic way in which the water flows, Justin only has to wait 3 time steps in order to figure out the initial water level in each pool. At time step $t = 0$, Justin can measure the initial water level in pool 1. At time $t = 1$, the water level in pool 1 will be equivalent to the initial water level in pool 2, and at time $t = 2$, the water level in pool 1 will be equivalent to the initial water level in pool 3. Thus Justin needs to wait 3 time steps (including the initial time step) – $t = 0, t = 1, and t = 2$ – in order to figure out the initial water levels in all the pools.

(b) Consider now a general pump matrix **A** that is known to Justin, not necessarily the one in the example above. Just for this part, suppose Justin had been told the initial water levels $\vec{x}[0]$ by someone else. Could he figure out $\vec{x}[t]$? **Write an expression for $\vec{x}[t]$ given the initial levels $\vec{x}[0]$ and the pump matrix A.**

Answer:

If Justin knows $\vec{x}[0]$, he can use what he knows about how the system changes over time (the *dynamics* of the system) to figure out $\vec{x}[t]$. These dynamics are summarized in the pump matrix **A**. **A** tells us how $\vec{x}$ updates at each time step. Every time we multiply on the left by **A**, we figure out what $\vec{x}$ will be one time step in the future. Thus we can calculate

$$\vec{x}[1] = \mathbf{A}\vec{x}[0]$$
$$\vec{x}[2] = \mathbf{A}\vec{x}[1] = \mathbf{A}(\mathbf{A}\vec{x}[0]) = \mathbf{A}^2\vec{x}[0]$$
$$\vec{x}[3] = \mathbf{A}\vec{x}[2] = \mathbf{A}^3\vec{x}[0])$$

Thus the general expression for $\vec{x}[t]$ in terms of **A** and $\vec{x}[0]$ is

$$\vec{x}[t] = \mathbf{A}^t\vec{x}[0]$$

This expression is useful for calculating the initial water levels (as we will see in part (d)). It is also interesting because it points out that if Justin is able to calculate $\vec{x}[0]$, then he really is able to calculate the water level of each pool at any time step in the future using this equation.

(c) Suppose we use $y[t]$ to denote Justin's measurement of the water level in pool 1 at time t. We know that $y[t] = x_1[t]$. **Find a vector $\vec{c}$ such that**

$$y[t] = \vec{c}^T \vec{x}[t]$$

Answer:

In this problem, we want to mathematically describe what Justin is doing when he measures the water level in his pool. He is measuring x_1 which is the first element of the vector $\vec{x}$. We can describe this measuring process as taking the multiplication of the specific vector $\vec{c}^T$ and $\vec{x}$. That vector $\vec{c}$ is given by

$$\vec{c} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Thus we have

$$y[t] = \vec{c}^T \vec{x}[t] = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1[t] \\ x_2[t] \\ x_3[t] \end{bmatrix} = 1 \cdot x_1[t] + 0 \cdot x_2[t] + 0 \cdot x_3[t] = x_1[t]$$

as desired.

(d) We want to know if tracking the water level in pool 1 is enough to eventually figure out the initial water level in all the pools. First **find a matrix D in terms of $\vec{c}$ and A (and powers of A), such that**

$$\begin{bmatrix} y[0] \\ y[1] \\ \vdots \\ y[T-1] \end{bmatrix} = \begin{bmatrix} & \\ & \mathbf{D} & \\ & \end{bmatrix} \begin{bmatrix} x_1[0] \\ x_2[0] \\ x_3[0] \end{bmatrix}$$

*(Hint: Think about what the rows of **D** should be. It suffices to give an expression for the jth row $\mathbf{D}_j$ of **D**.)*

Answer:

Now we start getting into the heart of the problem. We want to figure out how Justin can calculate the initial water levels even if the pump matrix **A** is not as simple as in part (a).

Over time Justin will acquire a bunch of measurements of the water level in his pool, $y[0], y[1], \ldots, y[t]$. We want to write these measurements in terms of $\vec{c}$, **A**, and $\vec{x}[0]$, so that we can write a system of equations that will allow us to solve for $\vec{x}[0]$.

Using the expressions from parts (b) and (c), we can write

$$y[t] = \vec{c}^T \vec{x}[t] = \vec{c}^T A^t \vec{x}[0]$$

Writing each measurement as a row of a matrix equation, we get

$$\begin{bmatrix} y[0] \\ y[1] \\ \vdots \\ y[T-1] \end{bmatrix} = \begin{bmatrix} \vec{c}^T \vec{x}[0] \\ \vec{c}^T \mathbf{A} \vec{x}[0] \\ \vdots \\ \vec{c}^T \mathbf{A}^{T-1} \vec{x}[0] \end{bmatrix} = \underbrace{\begin{bmatrix} \vec{c}^T \\ \vec{c}^T \mathbf{A} \\ \vdots \\ \vec{c}^T \mathbf{A}^{T-1} \end{bmatrix}}_{\mathbf{D}} \vec{x}[0]$$
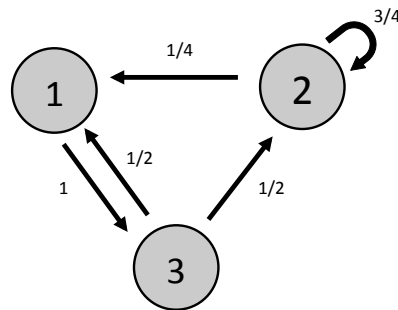
Thus if you count the first row as row 1, the second row as row 2, etc, then the *j*th row of **D**, $\mathbf{D}_j$, can be written as

$$\mathbf{D}_j = \vec{c}^T \mathbf{A}^{j-1}$$

It was also fine if you counted the first row as row 0, the second row as row 1, etc. In this case, D_j, is given by

$$\mathbf{D}_j = \vec{c}^T \mathbf{A}^j$$

(e) Now assume we have a specific network of pumps with a different pump matrix.

$$\begin{bmatrix} x_1[t+1] \\ x_2[t+1] \\ x_3[t+1] \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & \frac{1}{4} & \frac{1}{2} \\ 0 & \frac{3}{4} & \frac{1}{2} \\ 1 & 0 & 0 \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} x_1[t] \\ x_2[t] \\ x_3[t] \end{bmatrix}$$

Given this specific **A** matrix, **how many time steps T of observations in pool 1 will Justin need in order to recover the initial water levels $\vec{x}[0]$? Argue why this number of observations is enough.**

Answer:

Using the equation from part (d), we can start to see when Justin will be able to recover $\vec{x}[0]$ based on a specific **A** matrix. In order to find a unique $\vec{x}[0]$, Justin needs to be able to solve the equation

$$\begin{bmatrix} y[0] \\ y[1] \\ \vdots \\ y[T-1] \end{bmatrix} = \underbrace{\begin{bmatrix} \vec{c}^T \\ \vec{c}^T \mathbf{A} \\ \vdots \\ \vec{c}^T \mathbf{A}^{T-1} \end{bmatrix}}_{\mathbf{D}} \vec{x}[0] \tag{1}$$

The first thing we notice is that after a few time steps, Justin will have more equations than unknowns. Often times when there are more equations than unknowns the system has no solution. However, in this case since each physical measurement y (the water level in pool 1) is the result of a physical process (water being pumped between pools according to **A**) starting from some real initial water level, $\vec{x}[0]$, we know that that real initial water level, $\vec{x}[0]$ will be a solution. Thus we know that a solution exists.

What is not obvious immediately is whether our system of equations gives us enough information to solve for the real solution. It is possible that there will be multiple $\vec{x}[0]$'s that solve this equation. In this case, we won't be able to figure out the initial water levels.

To see how this might happen, suppose that all but two of the measurements summarized in the equation (1) are redundant, i.e. the row vectors $\vec{c}^T, \vec{c}^T \mathbf{A}, \ldots, \vec{c}^T \mathbf{A}^{T-1}$ are linearly dependent living in some two dimensional subspace of $\mathbb{R}^3$. In this case if, we row reduce D, we will end up with zeros in all the rows except 2. This means that we will have only 2 pivots and thus 1 free variable. Therefore, D has a null space that contains more than just $\vec{0}$ (**D** has a non-trivial null space).

This is problematic because multiple different initial water levels will give the same measurements. Consider the true initial water level $\vec{x}[0]$ and a vector that lives in the null space of **D**, $\vec{x}_{NS}[0]$. Note that through our measurement process, we won't be able to tell the difference between $\vec{x}[0]$ and $\vec{x}[0] + \vec{x}_{NS}[0]$ since both of them will give the same sequence of measurements $y[0], y[1], \ldots y[T-1]$.

$$\begin{bmatrix} \mathbf{D} \end{bmatrix} \left(\vec{x}[0] + \vec{x}_{NS}[0] \right) = \begin{bmatrix} \mathbf{D} \end{bmatrix} \vec{x}[0] + \underbrace{\begin{bmatrix} \mathbf{D} \end{bmatrix} \vec{x}_{NS}[0]}_{\vec{0}} = \begin{bmatrix} \mathbf{D} \end{bmatrix} \vec{x}[0]$$

Thus in order to determine $\vec{x}[0]$ uniquely we need **at least 3 rows of D to be linearly independent.** It turns out that the first three rows $\vec{c}^T, \vec{c}^T \mathbf{A}$, and $\vec{c}^T \mathbf{A}^2$ are linearly independent **thus Justin needs only 3 time steps ($t = 0, t = 1, t = 2$) to solve for $\vec{x}[0]$.** In order to get full credit on this part, you had to show that the first three rows are linearly independent. The easiest way to show this is by doing Gaussian elimination which we will do when we solve for the answer in part (f). [It is perfectly ok to point to a later problem part when solving a problem.] Note that you could also use three other linearly independent rows to solve for the solution but the first three rows are the easiest to calculate.

Here we see also why we have to be careful if **A** is more complicated. For a complicated **A**, it is difficult to tell if the vectors $\vec{c}^T, \vec{c}^T \mathbf{A}, \vec{c}^T \mathbf{A}^2, \ldots$ will span all of $\mathbb{R}^3$ until we calculate them out. If they don't

span $\mathbb{R}^3$, then **D** will have a non-trivial null space and this null space represents a set of initial water levels that are *hidden* from Justin's measurements. There is a general condition called *observability of* **A** *and* $\vec{c}$ that tells us when this null space exists. It has to do with the relationship between $\vec{c}$ and the eigenvectors of the matrix **A**. Maybe we'll get it to it later in the course or in EE16B.

(f) For the T chosen in the previous part and the pump matrix **A** given there, suppose Justin measures

$$y[t] = 1 \qquad \text{for } t = 0, 1, ..., (T-1)$$

What was $\vec{x}[0]$?

Answer:

In this part we will solve for $\vec{x}[0]$ using Gaussian elimination and in doing so we will show that the first three rows of **D** are linearly independent.

We want to solve the system of equations

$$\begin{bmatrix} y[0] \\ y[1] \\ y[2] \end{bmatrix} = \underbrace{\begin{bmatrix} \vec{c}^T \\ \vec{c}^T \mathbf{A} \\ \vec{c}^T \mathbf{A}^2 \end{bmatrix}}_{\bar{\mathbf{A}} = \begin{bmatrix} -\mathbf{A}_1- \\ -\mathbf{A}_2- \\ -\mathbf{A}_3- \end{bmatrix}} \begin{bmatrix} x_1[0] \\ x_2[0] \\ x_3[0] \end{bmatrix} \tag{2}$$

We first need to calculate $\vec{c}^T \mathbf{A}$ and $\vec{c}^T \mathbf{A}^2$.

$$\vec{c}^T \mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{4} & \frac{1}{2} \\ 0 & \frac{3}{4} & \frac{1}{2} \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{4} & \frac{1}{2} \end{bmatrix} \qquad \text{(This is just the 1st row of } \mathbf{A}\text{)}$$

$$\vec{c}^T \mathbf{A}^2 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{4} & \frac{1}{2} \\ 0 & \frac{3}{4} & \frac{1}{2} \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{4} & \frac{1}{2} \\ 0 & \frac{3}{4} & \frac{1}{2} \\ 1 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{3}{16} & \frac{1}{8} \\ \frac{1}{2} & \frac{9}{16} & \frac{3}{8} \\ 0 & \frac{1}{4} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{3}{16} & \frac{1}{8} \end{bmatrix} \qquad \text{(This is just the 1st row of } \mathbf{A}^2\text{)}$$

Writing out the augmented system for Equation (2) given the values $y[0] = y[1] = y[2] = 1$ we have

$$\begin{bmatrix} -\vec{c}^T- & \Big| & y[0] \\ -\vec{c}^T \mathbf{A}- & \Big| & y[1] \\ -\vec{c}^T \mathbf{A}^2- & \Big| & y[2] \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \Big| & 1 \\ 0 & \frac{1}{4} & \frac{1}{2} & \Big| & 1 \\ \frac{1}{2} & \frac{3}{16} & \frac{1}{8} & \Big| & 1 \end{bmatrix}$$

Row reducing, we get

$$\begin{bmatrix} 1 & 0 & 0 & \Big| & 1 \\ 0 & 1/4 & 1/2 & \Big| & 1 \\ 1/2 & 3/16 & 1/8 & \Big| & 1 \end{bmatrix} \underset{\substack{R_2 \leftarrow 4 \cdot R_2 \\ R_3 \leftarrow 16 \cdot R_3}}{\Longrightarrow} \begin{bmatrix} 1 & 0 & 0 & \Big| & 1 \\ 0 & 1 & 2 & \Big| & 4 \\ 8 & 3 & 2 & \Big| & 16 \end{bmatrix} \underset{R_3 \leftarrow R_3 - 8 \cdot R_1}{\Longrightarrow} \begin{bmatrix} 1 & 0 & 0 & \Big| & 1 \\ 0 & 1 & 2 & \Big| & 4 \\ 0 & 3 & 2 & \Big| & 8 \end{bmatrix}$$

$$\underset{R_3 \leftarrow R_3 - 3 \cdot R_2}{\Longrightarrow} \begin{bmatrix} 1 & 0 & 0 & \Big| & 1 \\ 0 & 1 & 2 & \Big| & 4 \\ 0 & 0 & -4 & \Big| & -4 \end{bmatrix} \underset{R_3 \leftarrow -4 \cdot R_3}{\Longrightarrow} \begin{bmatrix} 1 & 0 & 0 & \Big| & 1 \\ 0 & 1 & 2 & \Big| & 4 \\ 0 & 0 & 1 & \Big| & 1 \end{bmatrix} \underset{R_2 \leftarrow R_2 - 2 \cdot R_3}{\Longrightarrow} \begin{bmatrix} 1 & 0 & 0 & \Big| & 1 \\ 0 & 1 & 0 & \Big| & 2 \\ 0 & 0 & 1 & \Big| & 1 \end{bmatrix}$$

Thus we have that

$$\vec{x}[0] = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

We also have shown that $\bar{\mathbf{D}}$ row reduces to the identity. Thus the first 3 rows of $\mathbf{D}$ are linearly independent and $\vec{0}$ is the only vector in the null space of $\mathbf{D}$ and there is a unique solution for Equation (2).

Some other methods you could have used to solve for $\vec{x}[0]$ and show linear independence were solving for $\bar{\mathbf{D}}^{-1}$ by row reducing the augmented system

$$\left[\begin{array}{c:c} \mathbf{I} & \bar{\mathbf{D}} \end{array} \right] \implies \left[\begin{array}{c:c} \bar{\mathbf{D}}^{-1} & \mathbf{I} \end{array} \right] \qquad\qquad \vec{x}[0] = \left[\begin{array}{c} \bar{\mathbf{D}}^{-1} \end{array} \right] \begin{bmatrix} y[0] \\ y[1] \\ y[2] \end{bmatrix}$$

or using your initial knowledge that $x_1[0] = 1$ to write a system of 2 equations in the 2 unknowns, $x_2[0]$ and $x_3[0]$, and solve that system. (This is equivalent to starting at the third step of the row reduction shown above.)