EE263 Homework 1

Summer 2024

2.20. State equations for a linear mechanical system. The equations of motion of a lumped mechanical system undergoing small motions can be expressed as

$$M\ddot{q} + D\dot{q} + Kq = f$$

where $q(t) \in \mathbb{R}^k$ is the vector of deflections, M, D, and K are the *mass*, *damping*, and *stiffness* matrices, respectively, and $f(t) \in \mathbb{R}^k$ is the vector of externally applied forces. Assuming M is invertible, write linear system equations for the mechanical system, with state

$$x = \begin{bmatrix} q \\ \dot{q} \end{bmatrix},$$

input $u = f$, and output $y = q$.

Solution. We need to express the output q and the state derivative, $\dot{q}$ and $\ddot{q}$, as a linear function of the state variables q, $\dot{q}$ and the input f. In other words, we should find matrices A, B, C and D such that

$$\begin{bmatrix} \dot{q} \\ \ddot{q} \end{bmatrix} = A \begin{bmatrix} q \\ \dot{q} \end{bmatrix} + Bf, \quad q = C \begin{bmatrix} q \\ \dot{q} \end{bmatrix} + Df.$$

Matrices C and D are easy to find: simply, for the second equation to hold we should have

$$C = \begin{bmatrix} I & 0 \end{bmatrix}, \quad D = 0.$$

A and B are a bit harder to find. We will use the differential equation to express $\ddot{q}$ in terms of q, $\dot{q}$ and f. From the given dynamics equation $M\ddot{q} + D\dot{q} + Kq = f$, and assuming M is invertible, we get

$$\ddot{q} = -M^{-1}Kq - M^{-1}D\dot{q} + M^{-1}f,$$

which expresses $\ddot{q}$ in terms of q, $\dot{q}$, and f. Now we can write the linear dynamical system equations for the system. In block matrix notation we have

$$\frac{d}{dt}\begin{bmatrix} q \\ \dot{q} \end{bmatrix} = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}D \end{bmatrix}\begin{bmatrix} q \\ \dot{q} \end{bmatrix} + \begin{bmatrix} 0 \\ M^{-1} \end{bmatrix} u, \quad y = \begin{bmatrix} I & 0 \end{bmatrix}\begin{bmatrix} q \\ \dot{q} \end{bmatrix},$$

so the matrices in linear dynamical system description are:

$$A = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}D \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ M^{-1} \end{bmatrix}, \quad C = \begin{bmatrix} I & 0 \end{bmatrix}, \quad D = 0.$$

2.100. A mass subject to applied forces. Consider a unit mass subject to a time-varying force $f(t)$ for $0 \le t \le n$. Let the initial position and velocity of the mass both be zero. Suppose that the force has the form $f(t) = x_j$ for $j - 1 \le t < j$ and $j = 1, \ldots, n$. Let y_1 and y_2 denote, respectively, the position and velocity of the mass at time $t = n$.

a) Find the matrix $A \in \mathbb{R}^{2\times n}$ such that $y = Ax$.

b) For $n = 4$, find a sequence of input forces $x_1, \ldots, x_n$ that moves the mass to position 1 with velocity 0 at time n.

Solution. Let $p(t)$ and $v(t)$ denote, respectively, the position and velocity of the mass at time t.

a) The velocity is the integral of the applied force:

$$\begin{aligned}
v(t) &= v(0) + \int_0^t f(\tau)\,d\tau \\
&= v(0) + \sum_{j=1}^{\lfloor t \rfloor} \int_{j-1}^{j} f(\tau)\,d\tau + \int_{\lfloor t \rfloor}^{t} f(\tau)\,d\tau \\
&= v(0) + \sum_{j=1}^{\lfloor t \rfloor} \int_{j-1}^{j} x_j\,d\tau + \int_{\lfloor t \rfloor}^{t} x_{\lfloor t \rfloor+1}\,d\tau \\
&= v(0) + \sum_{j=1}^{\lfloor t \rfloor} (\tau x_j)|_{\tau=j-1}^{\tau=j} + (\tau x_{\lfloor t \rfloor+1})|_{\tau=\lfloor t \rfloor}^{\tau=t} \\
&= v(0) + \sum_{j=1}^{\lfloor t \rfloor} x_j + (t - \lfloor t \rfloor) x_{\lfloor t \rfloor+1}.
\end{aligned}$$

In particular, because the mass is initially at rest (that is, $v(0) = 0$), the final velocity is

$$y_2 = v(n) = \sum_{j=1}^{n} x_j.$$

Similarly, the position is the integral of the velocity:

$$\begin{aligned}
p(t) &= p(0) + \int_0^t v(\tau)\,d\tau \\
&= p(0) + \int_0^t (v(0) + (v(\tau) - v(0)))\,d\tau \\
&= p(0) + v(0)t + \int_0^t (v(\tau) - v(0))\,d\tau \\
&= p(0) + v(0)t + \sum_{j=1}^{\lfloor \tau \rfloor} \int_{j-1}^{j} (v(\tau) - v(0))\,d\tau + \int_{\lfloor t \rfloor}^{t} (v(\tau) - v(0))\,d\tau \\
&= p(0) + v(0)t + \sum_{j=1}^{\lfloor t \rfloor} \int_{j-1}^{j} \left(\sum_{k=1}^{\lfloor \tau \rfloor} x_k + (\tau - \lfloor \tau \rfloor) x_{\lfloor \tau \rfloor+1} \right) d\tau \\
&\quad + \int_{\lfloor t \rfloor}^{t} \left(\sum_{k=1}^{\lfloor \tau \rfloor} x_k + (\tau - \lfloor \tau \rfloor) x_{\lfloor \tau \rfloor+1} \right) d\tau \\
&= p(0) + v(0)t + \sum_{j=1}^{\lfloor t \rfloor} \int_{j-1}^{j} \left(\sum_{k=1}^{j-1} x_k + (\tau - (j-1)) x_j) \right) d\tau
\end{aligned}$$

$$
\begin{aligned}
&+ \int_{\lfloor t \rfloor}^{t} \left(\sum_{k=1}^{\lfloor t \rfloor} x_k + (\tau - \lfloor t \rfloor) x_{\lfloor t \rfloor + 1} \right) d\tau \\
&= p(0) + v(0)t + \sum_{k=1}^{\lfloor t \rfloor} \left(\sum_{k=1}^{j-1} \tau x_k + \frac{1}{2}(\tau - (j-1))^2 x_j \right) \Bigg|_{\tau=j-1}^{\tau=j} \\
&\quad + \left(\sum_{k=1}^{\lfloor t \rfloor} \tau x_k + \frac{1}{2}(\tau - \lfloor t \rfloor)^2 x_{\lfloor t \rfloor + 1} \right) \Bigg|_{\tau=\lfloor t \rfloor}^{\tau=t} \\
&= p(0) + v(0)t + \sum_{j=1}^{\lfloor t \rfloor} \left(\sum_{k=1}^{j-1} x_k + \frac{1}{2} x_j \right) + \left(\sum_{k=1}^{\lfloor t \rfloor} x_k + \frac{1}{2}(t - \lfloor t \rfloor)^2 x_{\lfloor t \rfloor + 1} \right) \\
&= p(0) + v(0)t + \sum_{j=1}^{\lfloor t \rfloor} \sum_{k=1}^{j-1} x_k + \sum_{j=1}^{\lfloor t \rfloor} \frac{1}{2} x_j + \sum_{k=1}^{\lfloor t \rfloor} x_k + \frac{1}{2}(t - \lfloor t \rfloor)^2 x_{\lfloor t \rfloor + 1} \\
&= p(0) + v(0)t + \sum_{k=1}^{\lfloor t \rfloor} \sum_{j=k+1}^{\lfloor t \rfloor} x_k + \sum_{k=1}^{\lfloor t \rfloor} \frac{1}{2} x_k + \sum_{k=1}^{\lfloor t \rfloor} x_k + \frac{1}{2}(t - \lfloor t \rfloor)^2 x_{\lfloor t \rfloor + 1} \\
&= p(0) + v(0)t + \sum_{k=1}^{\lfloor t \rfloor} (\lfloor t \rfloor - k) x_k + \sum_{k=1}^{\lfloor t \rfloor} \frac{1}{2} x_k + \sum_{k=1}^{\lfloor t \rfloor} x_k + \frac{1}{2}(t - \lfloor t \rfloor)^2 x_{\lfloor t \rfloor + 1} \\
&= p(0) + v(0)t + \sum_{k=1}^{\lfloor t \rfloor} \left((\lfloor t \rfloor - k) + \frac{1}{2} + (t - \lfloor t \rfloor) \right) x_k + \frac{1}{2}(t - \lfloor t \rfloor)^2 x_{\lfloor t \rfloor + 1} \\
&= p(0) + v(0)t + \sum_{k=1}^{\lfloor t \rfloor} \left(t - k + \frac{1}{2} \right) x_k + \frac{1}{2}(t - \lfloor t \rfloor)^2 x_{\lfloor t \rfloor + 1}.
\end{aligned}
$$

In particular, because the mass is initially at rest at the origin (that is, $p(0) = 0$ and $v(0) = 0$), the final position is

$$
y_1 = p(n) = \sum_{j=1}^{n} \left(n - k + \tfrac{1}{2} \right) x_j.
$$

Thus, we obtain the following system of linear equations:

$$
\begin{aligned}
y_1 &= \sum_{j=1}^{n} (n - j + \tfrac{1}{2}) x_j, \\
y_2 &= \sum_{j=1}^{n} x_j.
\end{aligned}
$$

Since A_{ij} gives the coefficient of x_j in our expression for y_i, we have that

$$
A_{1j} = n - j + \tfrac{1}{2} \qquad \text{and} \qquad A_{2j} = 1, \qquad j = 1, \ldots, n.
$$

More concretely, we have that

$$A = \begin{bmatrix} n - \frac{1}{2} & n - \frac{3}{2} & \cdots & \frac{3}{2} & \frac{1}{2} \\ 1 & 1 & \cdots & 1 & 1 \end{bmatrix}.$$

b) We want to solve the following system of linear equations:

$$\begin{bmatrix} \frac{7}{2} & \frac{5}{2} & \frac{3}{2} & \frac{1}{2} \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

This system is underdetermined, and has infinitely many solutions. Suppose we choose $x_2 = x_3 = 0$. Then, we are left with the system

$$\begin{bmatrix} \frac{7}{2} & \frac{1}{2} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

The second equation implies that $x_4 = -x_1$. Then, the first equation becomes

$$\tfrac{7}{2}x_1 + \tfrac{1}{2}x_4 = \tfrac{7}{2}x_1 - \tfrac{1}{2}x_1 = 3x_1 = 1.$$

Solving this equation, we find that $x_1 = \frac{1}{3}$. Substituting this value into our expression for x_4 gives $x_4 = -x_1 = -\frac{1}{3}$. Thus, one sequence of input forces that moves the mass to position 1 with velocity 0 at time n is

$$x = \begin{bmatrix} \frac{1}{3} \\ 0 \\ 0 \\ -\frac{1}{3} \end{bmatrix}.$$

2.130. Most common symbol in a given position. Consider (again) the following Markov language. We have an alphabet of n symbols $1, 2, \ldots, n$. A sentence is a finite sequence of symbols, $k_1, \ldots, k_m$ where $k_i \in \{1, \ldots, n\}$. A language or code consists of a set of sequences, which we will call the *allowable sequences*. A language is called *Markov* if the allowed sequences can be described by giving the allowable transitions between consecutive symbols. For each symbol we give a set of symbols which are allowed to follow the symbol. As a simple example, consider a Markov language with three symbols $1, 2, 3$. Symbol 1 can be followed by 1 or 3; symbol 2 must be followed by 3; and symbol 3 can be followed by 1 or 2. The sentence 1132313 is allowable (*i.e.*, in the language); the sentence 1132312 is not allowable (*i.e.*, not in the language). To describe the allowed symbol transitions we can define a matrix $A \in \mathbb{R}^{n \times n}$ by

$$A_{ij} = \begin{cases} 1 & \text{if symbol } i \text{ is allowed to follow symbol } j \\ 0 & \text{if symbol } i \text{ is not allowed to follow symbol } j \end{cases}.$$

There are five symbols $1, 2, 3, 4, 5$, and the following symbol transition rules:

- 1 must be followed by 2 or 3

- 2 must be followed by 2 or 5
- 3 must be followed by 1
- 4 must be followed by 4 or 2 or 5
- 5 must be followed by 1 or 3

Among all allowed sequences of length 10, find the most common value for the seventh symbol. In principle you could solve this problem by writing down all allowed sequences of length 10, and counting how many of these have symbol i as the seventh symbol, for $i = 1, \ldots 5$. (We're interested in the symbol for which this count is largest.) But we'd like you to use a smarter approach. Explain clearly how you solve the problem, as well as giving the specific answer. Hint: you may find the interpretation of A^k helpful.

Solution. To solve this problem you need to understand the interpretation of A^k, where A is the symbol transition matrix associated with the language, *i.e.*,

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}.$$

Specifically, the ij entry of A^k is the number of allowed sequences of length $k+1$ that start with symbol j and end with symbol i. In the homework problem, you added up all entries of A^9 to get the total number of sequences of length 10. The answer was 1079. To find the number of allowed sequences that contain the sumbol i in the seventh position, we proceed as follows. We are interested in allowed sequernces that have the form

$$x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6,\ i,\ x_8,\ x_9,\ x_{10}.$$

Such sequences consist of an allowed sequence of length 7, that ends in i,

$$x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6,\ i,$$

followed by an allowed sequence of length 4 that starts with i,

$$i,\ x_8,\ x_9,\ x_{10}.$$

(Well, technically, we only write down i once when we concatenate the two sequences.) Conversely, if we have any allowed sequence of length 7 that ends in i, and any allowed sequence of length 4 that begins with i, we can put them together to get an allowed sequence of length 10 that has i in its seventh position. It follows that the total number of sequences of length 10 with i in the seventh position is the product of the number of sequences of length 7 that end in i, and the number of sequences of length 4 that begin with i. Whew. First let's find the total number of sequences of length 7 that end in symbol i. By the interpretation given above, this is equal to the sum of the entries in the ith row of A^6. In a similar way we find that the total number of allowed sequences of length 4 that start with i is given by the sum of the ith

column of A^3. Carrying out the required computations, we find the numbers of sequences of length 10 with i in the seventh place are, for $i = 1, \ldots, 5$,

$$288 \quad 448 \quad 144 \quad 14 \quad 185.$$

The most common symbol is $i = 2$. It occurs 448 times out the total of 1079 (which does agree with our previous calculations, which is always good ...) Some simple matlab code that carries out the required computations is:

```
A = [0 0 1 0 1;
1 1 0 1 0;
1 0 0 0 1;
0 0 0 1 0;
0 1 0 1 0];
A_1 = A^6; seq_1 = sum(A_1');
A_2 = A^3; seq_2 = sum(A_2);
seq = seq_1 .* seq_2
[occurences, most_common] = max(seq); most_common
tot_seq = sum(seq);
```

We should mention a common error. A number of people asserted that to find the most common symbol in the seventh position of a string that has length 10, we only need to find the most common symbol among strings of length 7. In this particular example, it is true that the symbol $i = 2$ is also the most common seventh symbol among all strings of length 7. But in general this can be false. The analysis above shows that you are maximizing a product of two numbers over i; the abbreviated method just maximizes the first number. Another error was to try to make the problem into a Markov chain problem, by normalizing the transitions to make them into probabilities, and then finding the symbol with maximum likelihood in the seventh position. This is not only wrong in general, it often led to the wrong answer as well (depending on how the transitions were normalized).

2.160. Some matrices from signal processing. We consider $x \in \mathbb{R}^n$ as a signal, with x_i the (scalar) value of the signal at (discrete) time period i, for $i = 1, \ldots, n$. Below we describe several transformations of the signal x, that produce a new signal y (whose dimension varies). For each one, find a matrix A for which $y = Ax$.

a) $2\times$ *up-conversion with linear interpolation.* We take $y \in \mathbb{R}^{2n-1}$. For i odd, $y_i = x_{(i+1)/2}$. For i even, $y_i = (x_{i/2} + x_{i/2+1})/2$. Roughly speaking, this operation doubles the sample rate, inserting new samples in between the original ones using linear interpolation.

b) $2\times$ *down-sampling.* We assume here that n is even, and take $y \in \mathbb{R}^{n/2}$, with $y_i = x_{2i}$.

c) $2\times$ *down-sampling with averaging.* We assume here that n is even, and take $y \in \mathbb{R}^{n/2}$, with $y_i = (x_{2i-1} + x_{2i})/2$.

Solution.

a)

$$A_{\text{lin-int}} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & & 0 & 0 \\ 0 & 1 & 0 & 0 & & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & \cdots & 0 & 0 \\ \vdots & & & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

b)

$$A_{\text{down-sam}} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

c)

$$A_{\text{down-sam-avg}} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & \cdots & 0 & 0 \\ \vdots & & & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1/2 & 1/2 \end{bmatrix}$$

2.170. Affine functions. A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is called *affine* if for any $x,\ y \in \mathbb{R}^n$ and any $\alpha,\ \beta \in \mathbb{R}$ with $\alpha + \beta = 1$, we have

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y).$$

(Without the restriction $\alpha + \beta = 1$, this would be the definition of linearity.)

a) Suppose that $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Show that the function $f(x) = Ax + b$ is affine.

b) Now the converse: Show that any affine function f can be represented as $f(x) = Ax + b$, for some $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. (This representation is unique: for a given affine function f there is only one A and one b for which $f(x) = Ax + b$ for all x.)

Hint. Show that the function $g(x) = f(x) - f(0)$ is linear.

You can think of an affine function as a linear function, plus an offset. In some contexts, affine functions are (mistakenly, or informally) called linear, even though in general they are not. (Example: $y = mx + b$ is described as 'linear' in US high schools.)

Solution.

a) With $f(x) = Ax + b$, we have

$$\begin{aligned} f(\alpha x + \beta y) &= A(\alpha x + \beta y) + b \\ &= \alpha Ax + \alpha b + \beta Ay + (1-\alpha)b \\ &= \alpha(Ax + b) + \beta(Ay + b) \\ &= \alpha f(x) + \beta f(y), \end{aligned}$$

and thus f is affine.

b) Assume that f is affine; we'll show that $g(x) = f(x) - f(0)$ is linear. First we show that $g(\alpha x) = \alpha g(x)$ for any $x \in \mathbb{R}^n$ and any $\alpha \in \mathbb{R}$.

$$\begin{aligned} g(\alpha x) &= f(\alpha x) - f(0) \\ &= f(\alpha x + (1-\alpha)0) - f(0) \\ &= \alpha f(x) + (1-\alpha) f(0) - f(0) \\ &= \alpha(f(x) - f(0)) \\ &= \alpha g(x), \end{aligned}$$

where the third line follows from affineness of f (since $\alpha + \beta = 1$ when $\beta = 1 - \alpha$). To establish linearity of g, we must also show that $g(x + y) = g(x) + g(y)$. We do this as follows.

$$\begin{aligned} g(x+y) &= f(x+y) - f(0) \\ &= f\left((1/2)(2x) + (1/2)(2y)\right) - f(0) \\ &= (1/2)f(2x) + (1/2)f(2y) - f(0) \\ &= (1/2)(f(2x) - f(0)) + (1/2)(f(2y) - f(0)) \\ &= (1/2)g(2x) + (1/2)g(2y) \\ &= g(x) + g(y). \end{aligned}$$

The third line is by affineness of f. The last line uses the result above, *i.e.*, $g(\alpha z) = \alpha g(z)$ for any α and z. So now we know that g is linear. It follows that there is an $A \in \mathbb{R}^{m \times n}$ for which $g(x) = Ax$ for any x. Thus we have $f(x) = g(x) + f(0) = Ax + f(0)$. With $b = f(0)$, we see that $f(x) = Ax + b$.

You might be interested in a way to find A and b directly from the affine function f. This is done as follows. First we set $b = f(0)$. Then we set $a_i = f(e_i) - b$, for $i = 1, \ldots, n$, where e_i is the ith unit vector. Then we have $A = [a_1 \cdots a_n]$. So to find A and b, you need to evaluate f a total of $n + 1$ times. Thereafter, we can *predict* what $f(x)$ will be, for *any* x, using the form $f(x) = Ax + b$.