EE263 Homework 1

Fall 2023

2.50. Some linear functions associated with a convolution system. Suppose that u and y are scalar-valued discrete-time signals (*i.e.*, sequences) related via convolution:

$$y(k) = \sum_j h_j u(k-j), \quad k \in \mathbb{Z},$$

where $h_k \in \mathbb{R}$. You can assume that the convolution is *causal*, *i.e.*, $h_j = 0$ when $j < 0$.

a) *The input/output (Toeplitz) matrix.* Assume that $u(k) = 0$ for $k < 0$, and define

$$U = \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(N) \end{bmatrix}, \quad Y = \begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(N) \end{bmatrix}.$$

Thus U and Y are vectors that give the first $N+1$ values of the input and output signals, respectively. Find the matrix T such that $Y = TU$. The matrix T describes the linear mapping from (a chunk of) the input to (a chunk of) the output. T is called the input/output or Toeplitz matrix (of size $N+1$) associated with the convolution system.

b) *The Hankel matrix.* Now assume that $u(k) = 0$ for $k > 0$ or $k < -N$ and let

$$U = \begin{bmatrix} u(0) \\ u(-1) \\ \vdots \\ u(-N) \end{bmatrix}, \quad Y = \begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(N) \end{bmatrix}.$$

Here U gives the *past input* to the system, and Y gives (a chunk of) the resulting future output. Find the matrix H such that $Y = HU$. H is called the Hankel matrix (of size $N+1$) associated with the convolution system.

Solution.

a) *The input/output (Toeplitz) matrix.* Since $h_j = 0$ for $j < 0$ and $u(k) = 0$ for $k < 0$ we have

$$y(k) = \sum_{j=0}^{k} h_j u(k-j) = \sum_{j=0}^{k} h_{k-j} u(j),$$

so for $k = 0, 1, 2, \ldots, N$,

$$\begin{aligned} y(0) &= h_0 u(0) \\ y(1) &= h_1 u(0) + h_0 u(1) \\ y(2) &= h_2 u(0) + h_1 u(1) + h_0 u(2) \\ &\vdots \\ y(N) &= h_N u(0) + h_{N-1} u(1) + h_{N-2} u(2) + \cdots + h_0 u(N) \end{aligned}$$

These equations can be written in matrix form as

$$\underbrace{\begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} h_0 & 0 & 0 & \cdots & 0 \\ h_1 & h_0 & 0 & \cdots & 0 \\ h_2 & h_1 & h_0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \\ h_N & h_{N-1} & \cdots & h_1 & h_0 \end{bmatrix}}_{T} \underbrace{\begin{bmatrix} u(0) \\ u(1) \\ u(2) \\ \vdots \\ u(N) \end{bmatrix}}_{U}$$

(Note: the matrix T above is such that its (i, j)th entry depends only on the value of $i - j$. In other words, it is "constant along the diagonals." Such matrices are called Toeplitz matrices.)

b) *The Hankel matrix.* Since $u(k) = 0$ for $k > 0$ or $k < -N$

$$y(k) = \sum_{j=k}^{N+k} h_j u(k - j)$$

so for $k = 0, 1, 2, \ldots, N$

$$\begin{aligned} y(0) &= h_0 u(0) + h_1 u(-1) + \cdots + h_N u(-N) \\ y(1) &= h_1 u(0) + h_2 u(-1) + \cdots + h_{N+1} u(-N) \\ y(2) &= h_2 u(0) + h_3 u(-1) + \cdots + h_{N+2} u(-N) \\ &\vdots \\ y(N) &= h_N u(0) + h_{N+1} u(-1) + \cdots + h_{N+N} u(-N) \end{aligned}$$

and therefore

$$\underbrace{\begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} h_0 & h_1 & h_2 & \cdots & h_N \\ h_1 & h_2 & h_3 & \cdots & h_{N+1} \\ h_2 & h_3 & h_4 & \cdots & h_{N+2} \\ \vdots & & & & \vdots \\ h_N & h_{N+1} & h_{N+2} & \cdots & h_{N+N} \end{bmatrix}}_{H} \underbrace{\begin{bmatrix} u(0) \\ u(-1) \\ u(-2) \\ \vdots \\ u(-N) \end{bmatrix}}_{U}$$

(Note: matrices such as H above that are "constant along the antidiagonals" are called Hankel matrices.

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:

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

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:

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

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

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

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

2.170. 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.

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

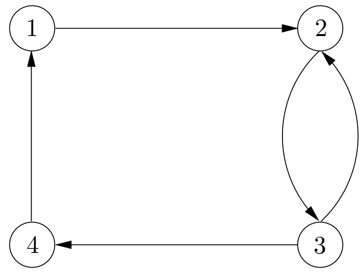
The third line is by affineness of f. The last line uses the result above, *i.e.*, $g(\alpha z) = \alpha g(z)$ for any α and z. So now we know that g is linear. It follows that there is an $A \in \mathbb{R}^{m\times n}$ for which $g(x) = Ax$ for any x. Thus we have $f(x) = g(x) + f(0) = Ax + f(0)$. With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

2.180. Paths and cycles in a directed graph. We consider a directed graph with n nodes. The graph is specified by its *node adjacency matrix* $A \in \mathbb{R}^{n\times n}$, defined as

$$A_{ij} = \begin{cases} 1 & \text{if there is an edge from node } j \text{ to node } i \\ 0 & \text{otherwise.} \end{cases}$$

Note that the edges are *oriented*, *i.e.*, $A_{34} = 1$ means there is an edge from node 4 to node 3. For simplicity we do not allow self-loops, *i.e.*, $A_{ii} = 0$ for all i, $1 \le i \le n$. A simple example illustrating this notation is shown below.



The node adjacency matrix for this example is

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

In this example, nodes 2 and 3 are connected in both directions, *i.e.*, there is an edge from 2 to 3 and also an edge from 3 to 2. A *path* of length $l > 0$ from node j to node i is a sequence $s_0 = j, s_1, \ldots, s_l = i$ of nodes, with $A_{s_{k+1},s_k} = 1$ for $k = 0, 1, \ldots, l-1$. For example, in the graph shown above, $1, 2, 3, 2$ is a path of length 3. A *cycle* of length l is a path of length l, with the same starting and ending node, with no repeated nodes other than the endpoints. In other words, a cycle is a sequence of nodes of the form $s_0, s_1, \ldots, s_{l-1}, s_0$, with

$$A_{s_1,s_0} = 1, \quad A_{s_2,s_1} = 1, \quad \ldots \quad A_{s_0,s_{l-1}} = 1,$$

and

$$s_i \neq s_j \text{ for } i \neq j, \quad i, j = 0, \ldots, l-1.$$

For example, in the graph shown above, $1, 2, 3, 4, 1$ is a cycle of length 4. The rest of this problem concerns a specific graph, given in the file `directed_graph.json` on the course web site. For each of the following questions, you must give the answer explicitly (for example, enclosed in a box). You must also explain clearly how you arrived at your answer.

a) What is the length of a shortest cycle? (Shortest means minimum length.)

b) What is the length of a shortest path from node 13 to node 17? (If there are no paths from node 13 to node 17, you can give the answer as 'infinity'.)

c) What is the length of a shortest path from node 13 to node 17, that *does not* pass through node 3?

d) What is the length of a shortest path from node 13 to node 17, that *does* pass through node 9?

e) Among all paths of length 10 that start at node 5, find the most common ending node.

f) Among all paths of length 10 that end at node 8, find the most common starting node.

g) Among all paths of length 10, find the most common pair of starting and ending nodes. In other words, find q, r which maximize the number of paths of length 10 from q to r.

Solution.

a) Recall that $(A^k)_{ij}$ gives the number of paths of length k from node j to node i. Thus, $(A^k)_{ii}$ is the number of paths of length k from node i to itself. Now imagine increasing k from $k = 1$ to $k = 2$, $k = 3$, and so on. We find the smallest k for which $(A^k)_{ii} > 0$. This k is the length of the smallest path from i to itself. This path is in fact also a cycle, since it cannot repeat nodes. (If it repeated nodes, there would have been a shorter path from i to itself.) Now let's solve the problem. To find the length of a shortest cycle, find the smallest k such that $(A^k)_{ii} > 0$ for some i. Note $k \leq n$, because if a cycle exists then it is at most of length n, where n is the number of nodes in the graph. The smallest cycle is of length 6.

b) To find the length of a shortest path from node 13 to node 17, find the smallest k such that $(A^k)_{17,13} > 0$. The shortest path from node 13 to node 17 is of length 4.

c) To find the shortest path from node 13 to node 17, that does not pass though node 3, remove node 3 from the graph and then find the shortest path from node 13 to node 17. The new adjacency matrix B for the graph is obtained by removing the 3rd row and column of the matrix A. Then find the smallest k such that $(B^k)_{17,13} > 0$. The shortest path from node 13 to node 17, that does not pass through node 3 is of length 5.

d) To find the smallest path from node 13 to node 17, that does pass through node 9, find the shortest path from node 13 to node 9 and the shortest path from node 9 to node 17. The shortest path from node 13 to node 17, that does pass through node 9 is of length 10.

e) The matrix A^{10} gives the number of paths of length 10, *i.e.*, $(A^{10})_{ij}$ is the number of paths of length 10 that start at node j and end at node i. The 5th column of the matrix A^{10} gives the number of paths of length 10 that start at node 5 and end at nodes $1, 2, \ldots, 20$ respectively. The index of the maximum entry of this column gives the most common ending node for paths of length 10 starting at node 5. The most common ending node for paths of length 10 starting at node 5, is 5.

f) The 8th row of the matrix A^{10} gives the number of paths that end at node 8. The index of the maximum entry of this row gives the most common starting node for paths of length 10 ending at node 8. The most common starting node for paths of length 10 ending at node 8, is 8.

g) The most common source/destination pair for paths of length 10 is the index of the maximum entry of A^{10}, *i.e.*, if (q, r) is the most common source/destination pair then no other number in the matrix A^{10} is greater than $(A^{10})_{rq}$. The most common source/destination pair for paths of length 10 is $(8, 17)$.

2.210. Express the following statements in matrix language. You can assume that all matrices mentioned have appropriate dimensions. Here is an example: "Every column of C is a linear combination of the columns of B" can be expressed as "$C = BF$ for some matrix F".

There can be several answers; one is good enough for us.

a) Suppose Z has n columns. For each i, row i of Z is a linear combination of rows $i, \ldots, n$ of Y.

b) W is obtained from V by permuting adjacent odd and even columns (*i.e.*, 1 and 2, 3 and 4, ...).

c) Each column of P makes an acute angle with each column of Q.

d) Each column of P makes an acute angle with the corresponding column of Q.

e) The first k columns of A are orthogonal to the remaining columns of A.

Solution.

a) $Z = UY$, where U is upper triangular, *i.e.*, $U_{ij} = 0$ for $i > j$.

b) $W = VS$, where S is the odd-even switch matrix, defined as

$$S = \begin{bmatrix} e_2 & e_1 & e_4 & e_3 & \cdots & e_m & e_{m-1} \end{bmatrix}.$$

c) All entries of the matrix $P^\mathsf{T} Q$ are positive.

d) The diagonal entries of the matrix $P^\mathsf{T} Q$ are positive.

e) $A^\mathsf{T} A$ has the block diagonal form

$$A^\mathsf{T} A = \begin{bmatrix} B_{11} & 0 \\ 0 & B_{22} \end{bmatrix},$$

where $B_{11} \in \mathbb{R}^{k \times k}$.

2.230. Population dynamics. An ecosystem consists of n species that interact (say, by eating other species, eating each other's food sources, eating each other's predators, and so on). We let $x(t) \in \mathbb{R}^n$ be the vector of deviations of the species populations (say, in thousands) from some equilibrium values (which don't matter here), in time period (say, month) t. In this model, time will take on the discrete values $t = 0, 1, 2, \ldots$. Thus $x_3(4) < 0$ means that the population of species 3 in time period 4 is below its equilibrium level. (It does not mean the population of species 3 is negative in time period 4.)

The population (deviations) follows a discrete-time linear dynamical system, which means that $x(t+1)$ is determined by $x(t)$. That is, we can compute the entire sequence $x(0), x(1), x(2), \ldots$ from $x(0)$ by applying the iteration

$$x(t+1) = Ax(t).$$

We refer to $x(0)$ as the *initial population perturbation*.

The questions below pertain to the specific case with $n = 10$ species, with matrix A given in `pop_dyn_data.json`.

a) Suppose the initial perturbation is $x(0) = e_4$ (meaning, we inject one thousand new creatures of species 4 into the ecosystem at $t = 0$). How long will it take to affect the other species populations? In other words, report a vector s, where s_i is the smallest t for which $x_i(t) \neq 0$. (We have $s_4 = 0$).

b) *Population control.* We can choose any initial perturbation that satisfies $|x_i(0)| \leq 1$ for each $i = 1, \ldots, 10$. (We achieve this by introducing additional creatures and/or hunting and fishing.) What initial perturbation $x(0)$ would you choose in order to maximize the population of species 1 at time $t = 10$? Explain your reasoning. Give the initial perturbation, and using your selected initial perturbation, give $x_1(10)$ and plot $x_1(t)$ versus t for $t = 0, \ldots, 40$.

Solution.

a) We can simply simulate the first few time periods by iterating $x(t+1) = Ax(t)$ and check when the elements of the vectors $x(t)$ first become nonzero. The code for this method is given below.

We find that $s = (4, 2, 1, 0, 4, 3, 3, 2, 4, 3)$.

Alternatively, since the system evolves by repeated application of the matrix A, $x(t) = A^t x(0)$. Since $x(0) = e_4$, we will have $x_i(t) \neq 0$ exactly when $(A^t)_{i,4} \neq 0$. We can thus solve this problem by examining the 4th columns of each of the first few powers (say, 10) of A, looking to see when the elements first become nonzero.

b) Let $B = A^{10}$, so $x(10) = A^{10}x(0) = Bx(0)$. We want to maximize $x_1(10)$, which is given by the inner product of the 1st row of B with $x(0)$. That is,

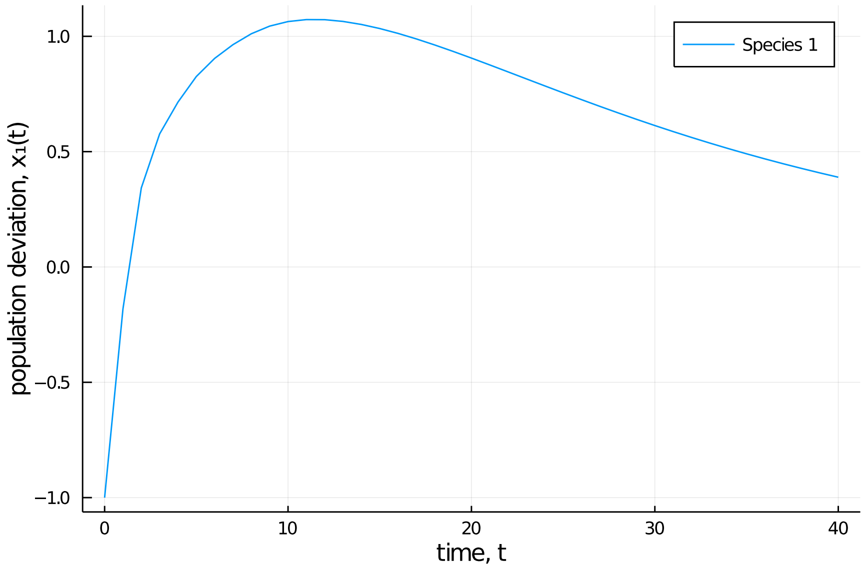
$$x_1(10) = \sum_{j=1}^{10} b_{1,j} x_j(0).$$

This sum is called *separable* because the jth term depends only on the variable $x_j(0)$. That means that the sum is maximized when each term is maximized separately. Since we are constrained to having the elements $|x_j(0)| \leq 1$, it's easy to see that the term $b_{1,j}x_j(0)$ is maximized when $x_j(0) = \text{sign}(b_{1,j})$, *i.e.*, $x_j(0)$ is either $+1$ or -1, with the sign matching the sign of $b_{1,j}$.

The code and the plot of $x_1(t)$ is given below. We find that

$$x(0) = (-1, -1, 1, -1, 1, 1, 1, 1, 1, -1)$$

and $x_1(10) = 1.06$. Note that, counterintuitively, the $x(0)$ which maximizes $x_1(10)$ sets the initial perturbation of species 1 to -1!




```
# cell 1
include("readclassjson.jl")
data = readclassjson("pop_dyn_data.json")
A = data["A"]
n = data["n"]

# cell 2
s = -ones(Int64, n)
x = zeros(n)
x[4] = 1.0
for t = 0:10
    s[(x .!= 0) .& (s .== -1)] .= t
    if all(s .!= -1)
        break
    end
    x = A*x
end
s

# cell 3
B = A^10
xat0 = sign.(B[1,:])

# cell 4
xat10 = B*xat0
xat10[1]
```

```
# cell 5
using Plots
x1history = zeros(41)
x = xat0
x1history[1] = x[1]
for t = 1:40
    x = A*x
    x1history[t+1] = x[1]
end
plot(0:40, x1history, label="Species 1",
    ylabel="population deviation, x1(t)", xlabel="time, t")
```