EECS 16A Designing Information Devices and Systems I
Summer 2020 Homework 2B

This homework is due Sunday July 12, 2020, at 23:59 PT.
Self-grades are due Wednesday July 15, 2020, at 23:59 PT.

Submission Format
Your homework submission should consist of a single PDF file that contains all of your answers (any hand-written answers should be scanned) as well as your IPython notebook saved as a PDF.

Please attach a PDF of your Jupyter notebook for all the problems that involve coding. Make sure the results of your plots (if any) are visible. Please assign the PDF of the notebook to the correct problems on Gradescope — we will be unable to grade the problems without this assignment or submission.

Homework Learning Goals: *This homework is all about applications for the linear algebra tools we have developed in the course so far.*

1. **Traffic Flows (PRACTICE)**

Learning Objective: *The learning objective of this problem is to see how the concept of nullspaces can be applied to flow problems.*

Your goal is to measure the flow rates of vehicles along roads in a town. It is prohibitively (too) expensive to place a traffic sensor along every road. You realize, however, that the number of cars flowing into an intersection must equal the number of cars flowing out. You can use this "flow conservation" to determine the traffic along all roads in a network by measuring the flow along only some roads. In this problem, we will explore this concept.

(a) Let's begin with a network with three intersections, A, B and C. Define the flow t_1 as the rate of cars (cars/hour) on the road between B and A, flow t_2 as the rate on the road between C and B, and flow t_3 as the rate on the road between C and A.

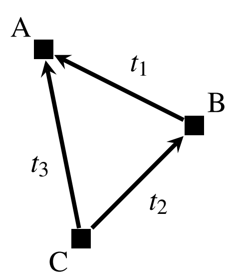


Figure 1: A simple road network.

(Note: The directions of the arrows in the figure are the way that we define positive flow by convention. For example, if there were 100 cars per hour traveling from A to C, then $t_3 = -100$. The flows now are not fractions of water in reservoirs as in the pumps setting, but numbers of cars.)

We assume the "flow conservation" constraints: the net number of cars per hour flowing into each intersection is zero. For example at intersection B, we have the constraint $t_2 - t_1 = 0$. The full set of constraints (one per intersection) is:

$$\begin{cases} t_1 + t_3 = 0 \\ t_2 - t_1 = 0 \\ -t_3 - t_2 = 0 \end{cases}$$

As mentioned earlier, we can place sensors on a road to measure the flow through it, but we have a limited budget, and we would like to determine all of the flows with the smallest possible number of sensors.

Suppose for the network above we have one sensor reading, $t_1 = 10$. Can we figure out the flows along the other roads? (That is, the values of t_2 and t_3). If we can, find the values of t_2 and t_3.

Solution:

Yes, since we know that $t_1 = t_2 = -t_3$, so we must have $t_2 = 10$ and $t_3 = -10$.

(b) Now suppose we have a larger network, as shown in Figure 2.

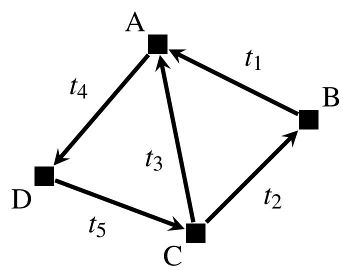


Figure 2: A larger road network.

We would again like to determine the traffic flows on all roads, using measurements from some sensors. A Berkeley student claims that we need two sensors placed on the roads CA (measuring t_3) and DC (measuring t_5). A Stanford student claims that we need two sensors placed on the roads CB (measuring t_2) and BA (measuring t_1). Write out the system of linear equations that represents this flow graph. Is it possible to determine all traffic flows, $\left[t_1, t_2, t_3, t_4, t_5\right]^T$, with the Berkeley student's suggestion? How about the Stanford student's suggestion?

Solution: Since we have 4 intersections, we can write 4 linear equations describing the flows into and out of each intersection. We know that the flows into and out of an intersection must sum to 0. The set of linear equations that represents this flow graph is:

$$\begin{cases} t_1 + t_3 - t_4 = 0 \\ t_2 - t_1 = 0 \\ t_5 - t_2 - t_3 = 0 \\ t_4 - t_5 = 0 \end{cases}$$

The Stanford student is wrong (obviously). Observing t_1 and t_2 is not sufficient, as t_3, t_4 and t_5 can still not be uniquely determined. Specifically, for any $\alpha \in \mathbb{R}$, the following flow satisfies the constraints and the measurements:

$$t_4 = \alpha$$
$$t_5 = \alpha$$
$$t_3 = \alpha - t_1$$

On the other hand, if we're given t_3 and t_5, we can uniquely solve for all the traffic densities as follows since we know the flow conservation constraints. From the set of linear equations we obtain:

$$\begin{aligned} t_1 &= t_5 - t_3 \\ t_2 &= t_5 - t_3 \\ t_4 &= t_5 \end{aligned}$$

This is related to the fact that t_3 and t_5 are parts of different loops in the graph, whereas t_1 and t_2 are in the same loop, so measuring both of them would not give additional information.

(c) We would like a more general way of determining the possible traffic flows in a network. Suppose we write the traffic flow on all roads as a vector $\vec{t} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \end{bmatrix}$. As a first step, let us try to write all the flow conservation constraints (one per intersection) as a matrix equation.
Construct a 4×5 matrix $\mathbf{B}$ such that the equation $\mathbf{B}\vec{t} = \vec{0}$:

$$\begin{bmatrix} & & \\ & \mathbf{B} & \\ & & \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

represents the flow conservation constraints for the network in Figure 2.
Hint: Each row is the constraint of an intersection. You can construct $\mathbf{B}$ using only $0, 1$, and -1 entries. This matrix is called the **incidence matrix**. What constraint does each column of $\mathbf{B}$ represent?
Solution:

$$\mathbf{B} = \begin{array}{c} \begin{bmatrix} 1 & 0 & 1 & -1 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} \begin{matrix} A \\ B \\ C \\ D \end{matrix} \\ \begin{matrix} t_1 & \;\; t_2 & \;\; t_3 & \;\; t_4 & \;\; t_5 \end{matrix} \end{array}$$

(The rows of this matrix can be in any order and your solution can differ by a factor of -1. However, the order of the elements within the row is still important and it must match the order of the elements of $\vec{t}$). Each row represents an intersection, and each column represents a road between two intersections. Each 1 on a row represents a road flowing into an intersection, and each -1 represents a road flowing out of an intersection. Each -1 in a column represents the source intersection of a road (where the arrow starts), and each 1 in a column represents the destination intersection of a road (where the arrow ends).
Each column of $\mathbf{B}$ must sum to 0. We expect each column to sum to 0 (and actually have exactly one -1 and one 1).

(d) Again, suppose we write the traffic flow on all roads as a vector $\vec{t} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \end{bmatrix}$. Then, determine the subspace of all valid traffic flows for the network of Figure 2. Notice that the set of all vectors $\vec{t}$ that satisfy $\mathbf{B}\vec{t} = \vec{0}$ is exactly the null space of the matrix $\mathbf{B}$. That is, we can find all valid traffic flows by computing the null space of $\mathbf{B}$. What is the dimension of the nullspace?

Solution:

We use Gaussian Elimination to find the nullspace.

$$\begin{bmatrix} 1 & 0 & 1 & -1 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} \underbrace{\Rightarrow}^{R_2+R_1\to R_2} \begin{bmatrix} 1 & 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & -1 & 0 \\ 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} \underbrace{\Rightarrow}^{R_2+R_3\to R_3} \begin{bmatrix} 1 & 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

$$\underbrace{\Rightarrow}^{R_4+R_5\to R_5} \begin{bmatrix} 1 & 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \underbrace{\Rightarrow}^{(-1)R_3\to R_3} \begin{bmatrix} 1 & 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \underbrace{\Rightarrow}^{R_2+R_3\to R_2} \begin{bmatrix} 1 & 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\underbrace{\Rightarrow}^{R_1+R_3\to R_1} \begin{bmatrix} 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

We see that we should let $t_3 = \alpha$ and $t_5 = \beta$, where α and β are free variables. The equations are:

$$\begin{aligned} t_1 &= \beta - \alpha \\ t_2 &= \beta - \alpha \\ t_3 &= \alpha \\ t_4 &= \beta \\ t_5 &= \beta \end{aligned}$$

The dimension of the nullspace is 2 because a minimum of 2 vectors are required to span the entire nullspace.

Note: We show here, for your reference, that the space of all possible traffic flows is a subspace. You don't not need to include this proof in your solution. Suppose we have a set of valid flows $\vec{t}$. Then, for any intersection, the total flow into it is the same as the total flow out of it. If we scale $\vec{t}$ by a constant a, each t_i will also get scaled by a. The total flows into and out of the intersection would be scaled by the same amount and remain equal to each other. Thus any scaling of a valid flow is still a valid flow. Suppose now we add valid flows $\vec{f}_1$ and $\vec{f}_2$ to get $\vec{t} = \vec{f}_1 + \vec{f}_2$. For any intersection I,

$$\text{total flow into } I = \text{total flow into } I \text{ from } \vec{f}_1 + \text{total flow into } I \text{ from } \vec{f}_2$$
$$\text{total flow out of } I = \text{total flow out of } I \text{ from } \vec{f}_1 + \text{total flow out of } I \text{ from } \vec{f}_2$$

Since the total flow into I from $\vec{f}_1$ is the same as the total flow out of I from $\vec{f}_1$ and similarly for $\vec{f}_2$, the total flow into I is the same as the total flow out of I. Therefore, the sum of any two valid flows is still a valid flow. Also, $\vec{t} = \vec{0}$ is a valid flow. Therefore the set of valid flows forms a subspace.

(e) Notice that we can represent the Berkeley student's measurement as $\mathbf{M}_B\vec{t}$, where:

$$\mathbf{M}_B\vec{t} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \vec{t} = \begin{bmatrix} t_3 \\ t_5 \end{bmatrix}$$

Write a matrix $\mathbf{M}_S$ that can be used to represent the Stanford student's measurement.

Solution:

$$\mathbf{M}_S\vec{t} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \vec{t} = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$$

(f) Now let us analyze more general road networks. Say there is a road network graph G, with incidence matrix $\mathbf{B}_G$. If $\mathbf{B}_G$ has a k-dimensional null space, does this mean measuring the flows along ***any*** k **roads** is always sufficient to recover all of the true flows? Prove or give a counterexample.

Hint: Consider the Stanford student from part (b).

Solution:

No, consider the network of Figure 2. The corresponding incidence matrix has a $k = 2$ dimensional null space, as you showed in part (e). However, measuring t_1 and t_2 (as the Stanford student suggested) is not sufficient, as you showed in part (b).

(g) (**Practice**) Assume that $\vec{u}$ and $\vec{t}$ are distinct valid flows, that is $\mathbf{B}_G\vec{u} = \mathbf{B}_G\vec{t} = \vec{0}$. Can you recover all of the network's true flows if $(\vec{u} - \vec{t})$ belongs to the nullspace of $\mathbf{M}_S$?

Clarification: A "valid" flow is one that is possible without violating the constraints on the nodes (so flow in must equal to flow out). There may be many valid flows, but only one "true" flow.

Solution: No. If $(\vec{u} - \vec{t})$ is in the nullspace of $\mathbf{M}_S$, it means $\mathbf{M}_S(\vec{u} - \vec{t}) = \vec{0}$. In other words, $\mathbf{M}_S\vec{u} = \mathbf{M}_S\vec{t}$. This means that two different flows will give us the same measurement, so the true flow cannot be recovered.

(h) (**Challenge: Practice**) If the incidence matrix $\mathbf{B}_G$ has a k-dimensional null space, does this mean we can **always pick a set of** k **roads** such that measuring the flows along these roads is sufficient to recover the exact flows? Prove or give a counterexample.

Solution:

Yes.

Let $\mathbf{U}$ be a matrix whose columns form a basis of the null space of $\mathbf{B}_G$, as above. The k columns of $\mathbf{U}$ are linearly independent since they form a basis. Since there are k linearly independent columns, when we run Gaussian elimination on $\mathbf{U}$, we must get k pivots. (Recall that "pivot" is the technical term for being able to row-reduce and turn a column into something that has exactly one 1 in it. The pivot is the entry that we found and turned into that 1.)

Therefore, the row space of $\mathbf{U}$ is k dimensional since there are some k linearly independent rows in $\mathbf{U}$ — namely the ones where we found pivots. Choose to measure the roads corresponding to these rows. This will work because:

For a given valid flow $\vec{t} = \mathbf{U}\vec{x}$, the results of measuring this flow vector are $\mathbf{U}^{(k)}\vec{x}$, where the matrix $\mathbf{U}^{(k)}$ is some k linearly independent rows of $\mathbf{U}$. By construction, the $k \times k$ matrix $\mathbf{U}^{(k)}$ has all linearly independent rows, so we can invert $\mathbf{U}^{(k)}$ to find $\vec{x}$ from $\mathbf{U}^{(k)}\vec{x}$ and then recover the flows along all the edges as $\mathbf{U}\vec{x}$.

This isn't the only set of k roads that will work. But it does provide a set of k roads that are guaranteed to work.

2. Noisy Images (PRACTICE)

Learning Goal: *The imaging lab uses the eigenvalues of the masking matrix to understand which masks are better than others for image reconstruction in the presence of additive noise. This problem explores the underlying mathematics.*

In lab, we used a single pixel camera to capture many measurements of an image $\vec{i}$. A single scalar measurement s_i is captured using a mask $\vec{h}_i$ such that $s_i = \vec{h}_i^T \vec{i}$. Many measurements can be expressed as a matrix-vector multiplication of the masks with the image, where the masks lie along the rows of the matrix.

$$\begin{bmatrix} s_1 \\ \vdots \\ s_N \end{bmatrix} = \begin{bmatrix} \vec{h}_1^T \\ \vdots \\ \vec{h}_N^T \end{bmatrix} \vec{i} \tag{1}$$

$$\vec{s} = \mathbf{H}\vec{i} \tag{2}$$

In the real world, noise, $\vec{w}$, creeps into our measurements, so instead,

$$\vec{s} = \mathbf{H}\vec{i} + \vec{w} \tag{3}$$

(a) Express $\vec{i}$ in terms of $\mathbf{H}$ (or its inverse), $\vec{s}$, and $\vec{w}$. Assume $\mathbf{H}$ is invertible. (*Hint:* Think about what you did in the imaging lab.)

Solution:

$$\vec{s} = \mathbf{H}\vec{i} + \vec{w} \tag{4}$$

$$\mathbf{H}^{-1}\vec{s} = \mathbf{H}^{-1}(\mathbf{H}\vec{i} + \vec{w}) \tag{5}$$

$$\mathbf{H}^{-1}\vec{s} = \mathbf{H}^{-1}\mathbf{H}\vec{i} + \mathbf{H}^{-1}\vec{w} \tag{6}$$

$$\mathbf{H}^{-1}\vec{s} = \vec{i} + \mathbf{H}^{-1}\vec{w} \tag{7}$$

$$\vec{i} = \mathbf{H}^{-1}\vec{s} - \mathbf{H}^{-1}\vec{w} \tag{8}$$

(b) Depending on how large or small the eigenvalues of $\mathbf{H}^{-1}$ are, we will amplify or attenuate our measurement's noise. The eigenvalues of $\mathbf{H}^{-1}$ are actually related to the eigenvalues of $\mathbf{H}$! Show that if λ is an eigenvalue of a matrix $\mathbf{H}$, then $\frac{1}{\lambda}$ is an eigenvalue of the matrix $\mathbf{H}^{-1}$.
Hint: Start with an eigenvalue λ and one corresponding eigenvector $\vec{v}$, such that they satisfy $\mathbf{H}\vec{v} = \lambda\vec{v}$.

Solution:

Since we're showing that $\frac{1}{\lambda}$ is an eigenvalue, we need to first show that $\lambda \neq 0$. We know that $\mathbf{H}$ is invertible,
$\Rightarrow \mathbf{H}\vec{x} = \vec{b}$ has a unique solution for all $\vec{b}$.
$\Rightarrow \mathbf{H}\vec{x} = \vec{0}$ has a unique solution.
$\Rightarrow \vec{x} = \vec{0}$ is the only solution to $\mathbf{H}\vec{x} = \vec{0}$.
$\Rightarrow \mathbf{H}\vec{x} = \vec{0}$ has no non-zero vectors $\vec{x}$ that satisfy it.

Therefore, 0 is not an eigenvalue. Let $\vec{v}$ be the eigenvector of $\mathbf{A}$ corresponding to λ.

$$\mathbf{H}\vec{v} = \lambda\vec{v}$$

Since we know that $\mathbf{H}$ is invertible, we can left-multiply both sides by $\mathbf{H}^{-1}$.

$$\mathbf{H}^{-1}\mathbf{H}\vec{v} = \lambda\mathbf{H}^{-1}\vec{v}$$

$$\vec{v} = \lambda\mathbf{H}^{-1}\vec{v}$$

$$\mathbf{H}^{-1}\vec{v} = \frac{1}{\lambda}\vec{v}$$

(c) We are going to try different $\mathbf{H}$ matrices in this problem and compare how they deal with noise. Run all of the cells in the attached IPython notebook. Which matrix performs best in reconstructing the original image and why? What do you observe regarding the eigenvalues of matrices $\mathbf{H_1}, \mathbf{H_2}$ and $\mathbf{H_3}$? What special matrix is $\mathbf{H}_1$? Notice that each plot in the iPython notebook returns the result of trying to image a noisy image as well as the minimum absolute value of the eigenvalue of each matrix. Comment on the effect of small eigenvalues on the noise in the image.

Solution:

See the associated `sol*.ipynb`, ipython file, where $\star$ is the number of this homework assignment.

Notice that we are printing the eigenvalue with the smallest absolute value. As the absolute value of the smallest eigenvalue of $\mathbf{H}$ decreases, the absolute value of the largest eigenvalue of $\mathbf{H^{-1}}$ increases (see why this happens in part d), hence the noise in the result increases. The matrix $\mathbf{H}_1$ is the identity matrix. Notice also that there are almost no visibile differences between the matrices $\mathbf{H}_2$ and $\mathbf{H}_3$.

(d) Now, because there is noise in our measurements, there will be noise in our recovered image. However, the noise is scaled. The noise in the recovered image, $\hat{\vec{w}}$, is related to $\vec{w}$, but it is transformed by $\mathbf{H}^{-1}$. Specifically,

$$\hat{\vec{w}} = \mathbf{H}^{-1}\vec{w} \tag{9}$$

To analyze how this transformation alters $\vec{w}$, consider representing $\vec{w}$ as a linear combination of the eigenvectors of $\mathbf{H}^{-1}$,

$$\vec{w} = \alpha_1\vec{b}_1 + \ldots + \alpha_N\vec{b}_N. \tag{10}$$

Where, $\vec{b}_i$ is $\mathbf{H}^{-1}$'s eigenvector corresponding to eigenvalue $\frac{1}{\lambda_i}$.

Show that we can express the recovered image's noise as,

$$\hat{\vec{w}} = \mathbf{H}^{-1}\vec{w} = \alpha_1\frac{1}{\lambda_1}\vec{b}_1 + \ldots + \alpha_N\frac{1}{\lambda_N}\vec{b}_N \tag{11}$$

Depending on the size of the eigenvalues, noise in the recovered image will be amplified or attenuated. For eigenvectors with large eigenvalues, will the noise signal along those eigenvectors be amplified of attenuated? For eigenvectors with small eigenvalues, will the noise signal along those eigenvectors be amplified of attenuated?

Solution: To show this, we will write $\vec{w}$ as a linear combination of the eigenvectors of $\mathbf{H^{-1}}$ and then use the distributivity property of matrix-vector multiplication operation:

$$\begin{aligned}\hat{\vec{w}} = \mathbf{H}^{-1}\vec{w} &= \mathbf{H}^{-1}(\alpha_1\vec{b}_1 + \ldots + \alpha_N\vec{b}_N) \\ &= \alpha_1\mathbf{H}^{-1}\vec{b}_1 + \ldots + \alpha_N\mathbf{H}^{-1}\vec{b}_N \\ &= \alpha_1\frac{1}{\lambda_1}\vec{b}_1 + \ldots + \alpha_N\frac{1}{\lambda_N}\vec{b}_N\end{aligned}$$

For eigenvectors with small eigenvalues, the noise signal will be amplified. This is bad and could corrupt the recovered image significantly.

For eigenvectors with large eigenvalues, the noise signal will be attenuated. This is better and will not corrupt the recovered image as much.

3. The Dynamics of Romeo and Juliet's Love Affair (PRACTICE)

In this problem, we will study a discrete-time model of the dynamics of Romeo and Juliet's love affair—adapted from Steven H. Strogatz's original paper, *Love Affairs and Differential Equations*, Mathematics Magazine, 61(1), p.35, 1988, which describes a continuous-time model.

Let $R[n]$ denote Romeo's feelings about Juliet on day n, and let $J[n]$ quantify Juliet's feelings about Romeo on day n. If $R[n] > 0$, it means that Romeo loves Juliet and inclines toward her, whereas if $R[n] < 0$, it means that Romeo is resentful of her and inclines away from her. A similar interpretation holds for $J[n]$, which represents Juliet's feelings about Romeo.

A larger $|R[n]|$ represents a more intense feeling of love (if $R[n] > 0$) or resentment (if $R[n] < 0$). If $R[n] = 0$, it means that Romeo has neutral feelings toward Juliet on day n. Similar interpretations hold for larger $|J[n]|$ and the case of $J[n] = 0$.

We model the dynamics of Romeo and Juliet's relationship using the following coupled system of linear evolutionary equations:

$$R[n+1] = a\,R[n] + b\,J[n], \quad n = 0,1,2,\ldots$$

and

$$J[n+1] = c\,R[n] + d\,J[n], \quad n = 0,1,2,\ldots,$$

which we can rewrite as

$$\vec{s}[n+1] = \mathbf{A}\,\vec{s}[n],$$

where

$$\vec{s}[n] = \begin{bmatrix} R[n] \\ J[n] \end{bmatrix}$$

denotes the state vector and

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

the state transition matrix for our dynamic system model.

The parameters a and d capture the linear fashion in which Romeo and Juliet respond to their own feelings, respectively, about the other person. It's reasonable to assume that $a, d > 0$, to avoid scenarios of fluctuating day-to-day mood swings. Within this positive range, if $0 < a < 1$, then the effect of Romeo's own feelings about Juliet tend to fizzle away with time (in the absence of influence from Juliet to the contrary), whereas if $a > 1$, Romeo's feelings about Juliet intensify with time (in the absence of influence from Juliet to the contrary). A similar interpretation holds when $0 < d < 1$ and $d > 1$.

The parameters b and c capture the linear fashion in which the other person's feelings influence $R[n]$ and $J[n]$, respectively. These parameters may or may not be positive. If $b > 0$, it means that the more Juliet shows affection for Romeo, the more he loves her and inclines toward her. If $b < 0$, it means that the more Juliet shows affection for Romeo, the more resentful he feels and the more he inclines away from her. A similar interpretation holds for the parameter c.

All in all, each of Romeo and Juliet has four romantic styles, which makes for a combined total of sixteen possible dynamic scenarios. The fate of their interactions depends on the romantic style each of them exhibits, the initial state, and the values of the entries in the state transition matrix $\mathbf{A}$. In this problem, we'll explore a subset of the possibilities.

(a) Consider the case where $a+b=c+d$ in the state-transition matrix

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

i. Show that

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

is an eigenvector of $\mathbf{A}$, and determine its corresponding eigenvalue λ_1. Also determine the other eigenpair $(\lambda_2, \vec{v}_2)$. Your expressions for λ_1, λ_2, and $\vec{v}_2$ must be in terms of one or more of the parameters a, b, c, and d.

Solution:

$$\begin{aligned} \mathbf{A}\begin{bmatrix} 1 \\ 1 \end{bmatrix} &= \begin{bmatrix} a+b \\ c+d \end{bmatrix} \\ &= (a+b)\begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= (c+d)\begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{aligned}$$

Let $\mu = a+b = c+d$. Then it's clear that $\begin{bmatrix} 1 & 1 \end{bmatrix}^T$ is an eigenvector of $\mathbf{A}$ corresponding to the eigenvalue μ. Therefore, the following is an eigenpair of $\mathbf{A}$:

$$\left(\lambda_1 = a+b = c+d, \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$$

To determine the other eigenpair $(\lambda_2, \vec{v}_2)$, we determine the other eigenvalue λ_2 first. We can do this in one of two ways:

Method I: Use the fact that $\text{tr}(\mathbf{A}) = \lambda_1 + \lambda_2$. We therefore have

$$\begin{aligned} a+d &= \lambda_1 + \lambda_2 \\ &= a+b+\lambda_2. \end{aligned}$$

Therefore,

$$\begin{aligned} \lambda_2 &= a+d-\lambda_1 \\ &= a+d-(a+b) \\ &= d-b. \end{aligned}$$

If we use the expression $\lambda_1 = c+d$, then an identical approach yields $\lambda_2 = a-c$.

Method II: An alternative approach is to determine the second eigenvalue λ_2 by solving the characteristic polynomial

$$\begin{aligned} \det(\lambda\mathbf{I} - \mathbf{A}) &= \det\left(\begin{bmatrix} \lambda - a & -b \\ -c & \lambda - d \end{bmatrix}\right) \\ &= (\lambda - a)(\lambda - d) - bc \\ &= \lambda^2 - (a+d)\lambda - bc \\ &= 0. \end{aligned}$$

We know from theory of quadratic polynomials that the sum of the roots equals the negative of the coefficient of the linear term λ. So, $\lambda_1 + \lambda_2 = a + d$. Notice that for a 2×2 matrix, the coefficient of λ is $-\text{tr}(\mathbf{A})$. And we can now use the same steps of Method I from here on.

Once we have the second eigenvalue, we use it to build the matrix $\lambda_2 \mathbf{I} - \mathbf{A}$. However, we do this in a smart way. We use the expression $\lambda_1 = a - c$ for the first row, and $\lambda_1 = d - b$ for the second row. That is,

$$\begin{aligned} \lambda_2 \mathbf{I} - \mathbf{A} &= \begin{bmatrix} (a-c)-a & -b \\ -c & (d-b)-d \end{bmatrix} \\ &= \begin{bmatrix} -c & -b \\ -c & -b \end{bmatrix}. \end{aligned}$$

Clearly, $\lambda_2 \mathbf{I} - \mathbf{A}$ has linearly dependent columns, and the vector

$$\vec{v}_2 = \begin{bmatrix} b \\ -c \end{bmatrix}$$

lies in its nullspace. Therefore, we have our second eigenpair:

$$\left(\lambda_2 = a - c = d - b, \vec{v}_2 = \begin{bmatrix} b \\ -c \end{bmatrix} \right).$$

Observation: You should note that any matrix whose row sums are a constant, say μ, must have $(\mu, \vec{1})$ as an eigenpair, where $\vec{1}$ is the all-ones vector of appropriate size.

ii. Consider the following state-transition matrix:

$$\mathbf{A} = \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix}$$

i. Determine the eigenpairs for this system.

Solution:

Notice that in this matrix, $a = d = 0.75$ and $b = c = 0.25$. So $\mu = a - c = d - b = 0.5$. Clearly, this is a row-stochastic matrix—each of its rows sums to 1. From the results of part (a)(i), we know that the eigenpairs of this matrix are

$$\left(\lambda_1 = 1, \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) \qquad \text{and} \qquad \left(\lambda_2 = 0.5, \vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right).$$

Observation: Notice that the eigenvectors $\vec{v}_1$ and $\vec{v}_2$ are orthogonal. This is not a coincidence. It turns out that the eigenvectors of a symmetric matrix are mutually orthogonal.

ii. Determine all the *fixed points* of the system. That is, find the set of points such that if Romeo and Juliet start at, or enter, any of those points, they'll stay in place forever: $\{\vec{s}_* \mid \mathbf{A}\vec{s}_* = \vec{s}_*\}$. Show these points on a diagram where the x and y-axes are $R[n]$ and $J[n]$.

Solution:

Any point along vector $\vec{s}_* = v_1 = \vec{1}$ is a fixed point, because $\vec{v}_1 = \vec{1}$ corresponds to the eigenvalue $\lambda_1 = 1$.

iii. Determine representative points along the state trajectory $\vec{s}[n]$, $n = 0, 1, 2, \ldots$, if Romeo and Juliet start from the initial state

$$\vec{s}[0] = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Solution:
The general solution is given by:

$$\begin{aligned}\vec{s}[n] &= \alpha_1 \lambda_1^n \vec{v}_1 + \alpha_2 \lambda_2^n \vec{v}_2 \\ &= \alpha_1 1^n \begin{bmatrix}1\\1\end{bmatrix} + \alpha_2 0.5^n \begin{bmatrix}1\\-1\end{bmatrix}\end{aligned}$$

Since $\vec{v}_1 \perp \vec{v}_2$ and since $\vec{s}[0] = \vec{v}_2$, we know that $\alpha_1 = 0$ and $\alpha_2 = 1$. Therefore,

$$\vec{s}[n] = 0.5^n \begin{bmatrix}1\\-1\end{bmatrix}.$$

Since 0.5^n decays to zero as $n \to \infty$, the state trajectory stays along the second eigenvector and decays to the origin:

$$\lim_{n\to\infty} (R[n], J[n]) = (0,0)$$

In particular, the state vector obeys the following trajectory:

$$\begin{bmatrix}R[n]\\J[n]\end{bmatrix} = \begin{bmatrix}\left(\frac{1}{2}\right)^n\\-\left(\frac{1}{2}\right)^n\end{bmatrix}, \quad n = 0,1,2,\ldots$$

This means that, ultimately, Romeo and Juliet will become neutral to each other.

iv. Suppose the initial state is $\vec{s}[0] = \begin{bmatrix}3 & 5\end{bmatrix}^T$. Determine a reasonably simple expression for the state vector $\vec{s}[n]$. Find the limiting state vector

$$\lim_{n\to\infty} \vec{s}[n].$$

Solution:
We must express the initial state vector as a linear combination of the eigenvectors. That is, we must solve the system of linear equations

$$\begin{gathered}\begin{bmatrix}\vec{v}_1 & \vec{v}_2\end{bmatrix} \vec{\alpha} = \vec{s}[0] \\ \begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}\begin{bmatrix}\alpha_1\\\alpha_2\end{bmatrix} = \begin{bmatrix}3\\5\end{bmatrix}.\end{gathered}$$

It's straightforward to find the solution:

$$\vec{\alpha} = \begin{bmatrix}\alpha_1\\\alpha_2\end{bmatrix} = \begin{bmatrix}4\\-1\end{bmatrix}.$$

Therefore, the state vector is given by

$$\begin{aligned}\vec{s}[n] &= \alpha_1 \lambda_1^n \vec{v}_1 + \alpha_2 \lambda_2^n \vec{v}_2 \\ &= 4\begin{bmatrix}1\\1\end{bmatrix} - \left(\frac{1}{2}\right)^n \begin{bmatrix}1\\-1\end{bmatrix} \\ &= \begin{bmatrix}4 - \left(\frac{1}{2}\right)^n \\ 4 + \left(\frac{1}{2}\right)^n\end{bmatrix}\end{aligned}$$

Clearly,

$$\lim_{n\to\infty} \vec{s}[n] = \begin{bmatrix}4\\4\end{bmatrix}.$$

(b) Consider the setup in which

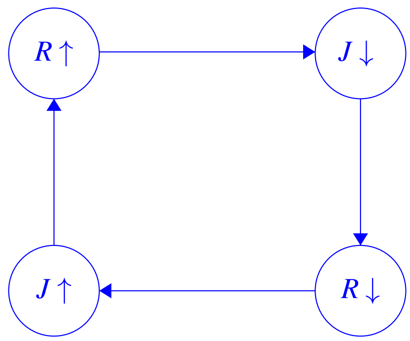
$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

In this scenario, if Juliet shows affection toward Romeo, Romeo's love for her increases, and he inclines toward her. The more intensely Romeo inclines toward her, the more Juliet distances herself. The more Juliet withdraws, the more Romeo is discouraged and retreats into his cave. But the more Romeo inclines away, the more Juliet finds him attractive and the more intensely she conveys her affection toward him. Juliet's increasing warmth increases Romeo's interest in her, which prompts him to incline toward her—again!

Predict the outcome of this scenario before you write down a single equation.

Solution:

We expect a never-ending cycle—an oscillation. The following diagram shows a qualitative picture of what happens.



Beginning with the top left node, we see that Romeo's affection increases. As a result, Juliet retreats, as depicted by the node on the top-right. In turn, this causes Romeo to lose hope and retreat, as shown in the bottom-right node. When Romeo pulls away, Juliet finds him mystically attractive and gravitates toward him, as shown by the bottom-left node. This causes Romeo to turn toward Juliet, which takes us back to the top-left node again, for yet another cycle.

Then determine a complete solution $\vec{s}[n]$ in the simplest of terms, assuming an initial state given by $\vec{s}[0] = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$. As part of this, you must determine the eigenvalues and eigenvectors of the $\mathbf{A}$.

Solution:

The eigenvalues are the roots of the equation

$$\begin{aligned} \det(\lambda \mathbf{I} - \mathbf{A}) &= \det \begin{bmatrix} \lambda & -1 \\ 1 & \lambda \end{bmatrix} \\ &= \lambda^2 + 1 = 0. \end{aligned}$$

So, $\lambda_1 = i$ and $\lambda_2 = -i$. Constructing the matrices $\lambda_1 \mathbf{I} - \mathbf{A}$ and $\lambda_2 \mathbf{I} - \mathbf{A}$, we find the corresponding eigenvectors by inspection:

$$\lambda_1 \mathbf{I} - \mathbf{A} = \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \quad \Longrightarrow \quad \vec{v}_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}$$

and

$$\lambda_2 \mathbf{I} - \mathbf{A} = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \quad \Longrightarrow \quad \vec{v}_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix}.$$

The matrix $\mathbf{A}$ has complex-valued eigenvalues and eigenvectors. Specifically, it has purely imaginary eigenvalues. This is not a coincidence. It turns out that if a matrix $\mathbf{A}$ has odd symmetry—that is, if $\mathbf{A}^T = -\mathbf{A}$—then its eigenvalues are purely imaginary.

Before we determine the general solution $\vec{s}[n]$, we must decompose the initial-state vector in terms of the two eigenvectors. The equation is

$$\underbrace{\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}}_{[\vec{v}_1 \;\; \vec{v}_2]} \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}}_{\vec{\alpha}} = \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\vec{s}[0]},$$

which yields the coefficient vector

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}.$$

The general solution is given by

$$\begin{aligned} \vec{s}[n] &= \alpha_1 \lambda_1^n \vec{v}_1 + \alpha_2 \lambda_2^n \vec{v}_2 \\ &= \frac{1}{2} i^n \begin{bmatrix} 1 \\ i \end{bmatrix} + \frac{1}{2} (-i)^n \begin{bmatrix} 1 \\ -i \end{bmatrix}. \end{aligned}$$

Since the two terms on the right-hand side are complex conjugates of one another, we have

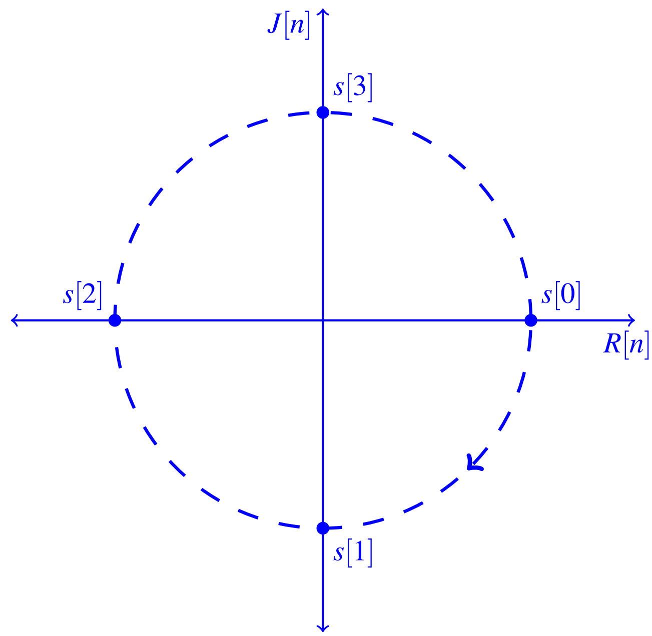
$$\begin{aligned} \vec{s}[n] &= 2 \,\mathrm{Re} \left\{ \frac{1}{2} i^n \begin{bmatrix} 1 \\ i \end{bmatrix} \right\} \\ &= \mathrm{Re} \left\{ \begin{bmatrix} i^n \\ i^{n+1} \end{bmatrix} \right\} \end{aligned}$$

$$\vec{s}[n] = \begin{cases} \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \text{if } n \geq 0 \text{ and } n \bmod 4 = 0 \\ \begin{bmatrix} 0 \\ -1 \end{bmatrix} & \text{if } n \geq 0 \text{ and } n \bmod 4 = 1 \\ \begin{bmatrix} -1 \\ 0 \end{bmatrix} & \text{if } n \geq 0 \text{ and } n \bmod 4 = 2 \\ \begin{bmatrix} 0 \\ 1 \end{bmatrix} & \text{if } n \geq 0 \text{ and } n \bmod 4 = 3 \end{cases}$$

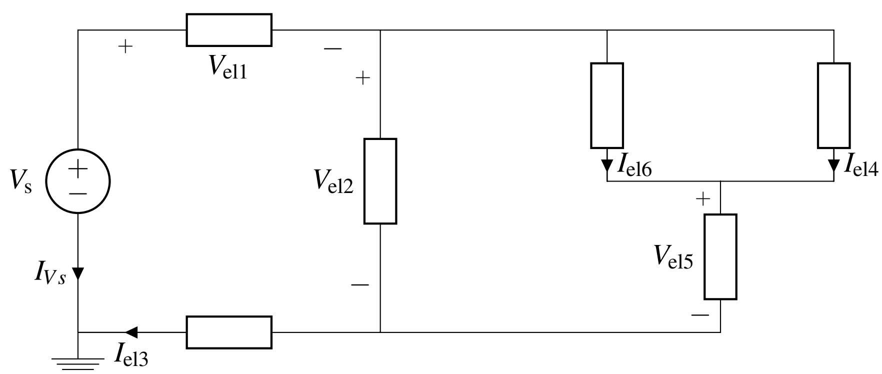
Plot (by hand, or otherwise without the assistance of any scientific computing software package), on a two-dimensional plane (called a *phase plane*)—where the horizontal axis denotes $R[n]$ and the vertical axis denotes $J[n]$—representative points along the trajectory of the state vector $\vec{s}[n]$, starting from the initial state given in this part. Describe, in plain words, what Romeo and Juliet are doing in this scenario. In other words, what does their state trajectory look like? Determine $\|\vec{s}[n]\|^2$ for all $n = 0, 1, 2, \ldots$ to corroborate your description of the state trajectory.

Solution:

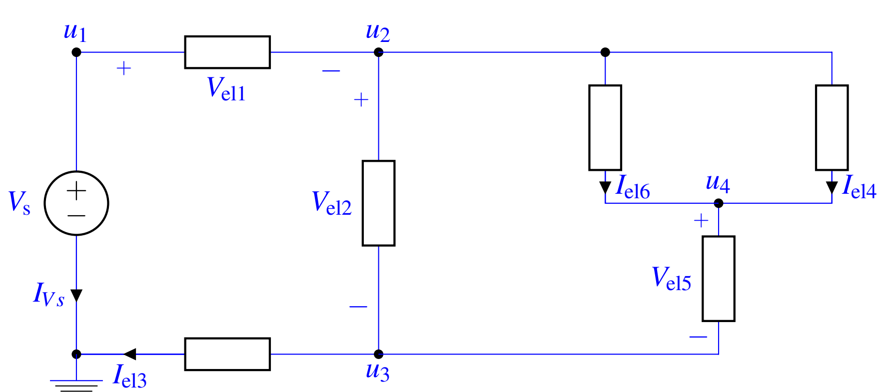
Romeo and Juliet are going around in a clockwise circle. Note that $\|\vec{s}[n]\|^2 = 1$ for all $n = 0, 1, 2, 3, \ldots$.



4. Intro to Circuits (MANDATORY - Not in scope for Midterm 1)

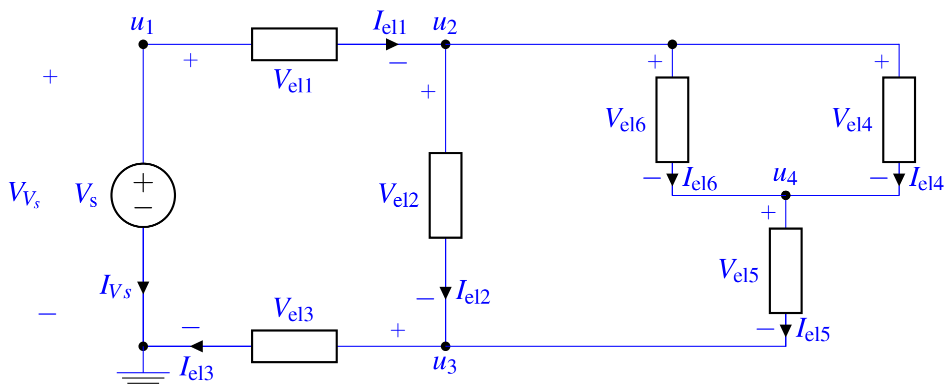


(a) How many nodes does the above circuit have? Label them.
Note: The ground node has been selected for you, so you don't need to label that, but you need to include it in your node count.
Solution: There is a total of 5 nodes in the circuit, including the ground node. They are labeled u_1 - u_4 below:



(b) Notice that elements 1 - 6 and the voltage source V_s have either the *voltage across* or the *current through* them not labeled. Label the missing *voltages across* or *currents through* for elements 1 - 6, and the voltage source V_s, so that they all follow **passive sign convention**.

Solution: The passive sign convention dictates that the current flows from the positive to the negative terminal of the element (or equivalently exiting the negative terminal / entering the positive terminal if you prefer):



(c) Express all element voltages (including the element voltage across the source, V_s) as a function of node voltages. This will be specific to the node labeling you chose in part (a).

Solution: For our specific node labeling we can write:

$$V_{V_s} = u_1 - 0 = u_1 (= V_s)$$

$$V_{\text{el1}} = u_1 - u_2$$

$$V_{\text{el2}} = u_2 - u_3$$

$$V_{\text{el3}} = u_3 - 0 = u_3$$

$$V_{\text{el4}} = u_2 - u_4$$

$$V_{\text{el5}} = u_4 - u_3$$

$$V_{\text{el6}} = u_2 - u_4$$

Notice that the element voltage is always of the form: $V_{\text{el}} = u_+ - u_-$.

(d) Write one KCL equation that involves the currents of elements 1 and 2.
Hint: This will **not** be specific to your node labeling.

Solution: The only node for which we can write a KCL involving elements 1 and 2 is node u_2, since they only intersect on that node:

$$I_{el1} = I_{el2} + I_{el6} + I_{el4}$$

(e) Write a KVL equation for all the loops that contain the voltage source V_s. These equations should be a function of element voltages and the voltage source V_s.
Hint: This will also **not** be specific to your node labeling. There are 3 such loops in the circuit.

Solution: Notice that there are in fact 3 loops that contain the voltage source V_s, for which we can write the following equations, starting each time from the ground node and ending at the ground node:

$$V_s - V_{\text{el1}} - V_{\text{el2}} - V_{\text{el3}} = 0$$
$$V_s - V_{\text{el1}} - V_{\text{el6}} - V_{\text{el5}} - V_{\text{el3}} = 0$$
$$V_s - V_{\text{el1}} - V_{\text{el4}} - V_{\text{el5}} - V_{\text{el3}} = 0$$

The reason this is not specific to our labeling is that the polarity of all elements is either given or set through the passive sign convention.

5. Homework Process and Study Group

Who else did you work with on this homework? List names and student ID's. (In case of homework party, you can also just describe the group.) How did you work on this homework?

Solution:

I worked on this homework with...

I first worked by myself for 2 hours, but got stuck on problem 5, so I went to office hours on...

Then I went to homework party for a few hours, where I finished the homework.