This homework is due February 6, 2017, at 23:59.
Self-grades are due February 9, 2017, at 23:59.

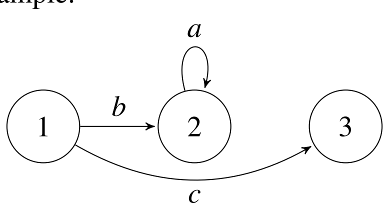
Submission Format
Your homework submission should consist of **two** files.

- `hw2.pdf`: A single pdf file that contains all your answers (any handwritten answers should be scanned) as well as your IPython notebook saved as a pdf. You can do this by printing the IPython notebook page in your browser and selecting the save to pdf option. Make sure any plots and results are showing. Also make sure you combine any separate pdfs into one file.
- `hw2.ipynb`: A single IPython notebook with all your code in it.

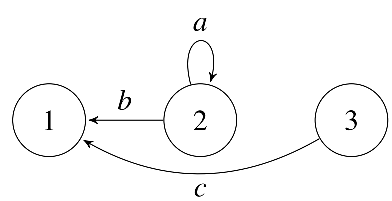
Submit each file to its respective assignment in Gradescope.

1. Transition Matrix Proofs

(a) Suppose there exists some network of websites, such as the "Original" example below. Assume the state vector at some time n is known. Would reversing the arrow directions, as shown in the "Reversed" example below, allow you to find the state vector at time $n-1$? If yes, argue why. If no, provide a counterexample.

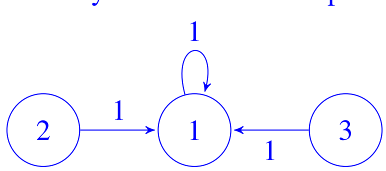


Original

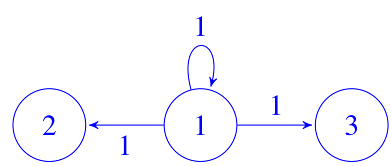


Reverse

Solution: Reversing the directions of the arrow is insufficient. Below is one counterexample. Note that there are many valid counterexamples other than the one shown here, all of which receive credit.



Original



Reverse

The original transition matrix is: $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. The reversed transition matrix is: $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$.

It is sufficient to show that a counterexample where the "original" state and the "reversed" state are not the same. For example:

$$\begin{bmatrix}1&1&1\\0&0&0\\0&0&0\end{bmatrix}\begin{bmatrix}1\\0.5\\0\end{bmatrix}=\begin{bmatrix}1.5\\0\\0\end{bmatrix} \qquad \begin{bmatrix}1&0&0\\1&0&0\\1&0&0\end{bmatrix}\begin{bmatrix}1.5\\0\\0\end{bmatrix}=\begin{bmatrix}1.5\\1.5\\1.5\end{bmatrix}$$

The original vectors $\begin{bmatrix}1\\0.5\\0\end{bmatrix}$ and the reversed vector $\begin{bmatrix}1.5\\1.5\\1.5\end{bmatrix}$ are not the same.

Note that "reversing the arrows" is equivalent to taking the transpose of our transition matrix. Physically, reversing the arrows equates to all of the "out" edges to each website becoming "in" edges; thus the matrix's columns become the rows, forming the transpose. In general, the transpose of a matrix is not the inverse. Note, the reversed system will not always physically make sense.

Alternate solution: Another way to show this is by multiplying the matrices. Recall that the product of a matrix and its inverse is $\mathbf{A}\mathbf{A}^{-1}=\mathbf{A}^{-1}\mathbf{A}=\mathbf{I}$. To uniquely go back to a previous state, the inverse must exist. Going back to our counterexample: if the transpose was the inverse, multiplying them would yield the identity, which does not occur:

$$\begin{bmatrix}1&1&1\\0&0&0\\0&0&0\end{bmatrix}\begin{bmatrix}1&0&0\\1&0&0\\1&0&0\end{bmatrix}=\begin{bmatrix}3&0&0\\0&0&0\\0&0&0\end{bmatrix}$$

Additionally, consider the example we provided at the beginning of the problem:

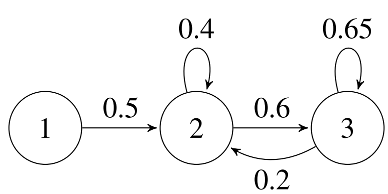
$$\begin{bmatrix}0&0&0\\b&a&0\\c&0&0\end{bmatrix}\begin{bmatrix}0&b&c\\0&a&0\\0&0&0\end{bmatrix}=\begin{bmatrix}0&0&0\\0&b^2+a^2&bc\\0&bc&c^2\end{bmatrix}$$

Since both b and c are nonzero terms, bc is a nonzero term. Thus this is not an identity matrix, so the transpose is not the inverse. (You can also note that the top left value is 0, but should be 1 for an identity).

(b) Suppose there is a state transition matrix such that the entries of each column vector sum to one. What is the physical interpretation about the total amount of people in the system?

Solution: The total amount of people in the system remains the same. There are no "leaks/drains" or "inlets/gains" in the system.

(c) Set up the state transition matrix $\mathbf{A}$ for the network shown below. Explain what this $\mathbf{A}$ matrix implies physically about the total amount of people in this system. (Note: If there is no "self arrow / self loop," then the people do not return to the original website.)



Solution:

$$\begin{bmatrix}0&0&0\\0.5&0.4&0.2\\0&0.6&0.65\end{bmatrix}$$

Notice that the entries in the column vectors do *not* sum to one. This is physically interpreted as a "leak" - i.e. the total amount of people is not conserved.

Note: This problem was adjusted. If you correctly described what happens to the amount of people in websites 1 and 3, then you will receive full credit. The amount of people in website 1 becomes zero after one time step, whereas the number of people in website 3 keeps decreasing.

(d) There is a state transition matrix where the entries of its rows sum to one. Prove that applying this system to a uniform vector will return the same uniform vector. A uniform vector is a vector where all the elements are the same.

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & a_{ij} & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x \\ \vdots \\ x \end{bmatrix} = \begin{bmatrix} x \\ \vdots \\ x \end{bmatrix}$$

Solution: Consider the row-interpretation of matrix multiplication: Each b_i is equal to the dot product of the row $\vec{\alpha}_i{}^T$ and the vector $\vec{x}$.

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & a_{ij} & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \vec{\alpha}_1{}^T \\ \vec{\alpha}_2{}^T \\ \vdots \\ \vec{\alpha}_n{}^T \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Let's look at the example for the first row b_i.

$$b_1 = \vec{\alpha}_1{}^T \vec{x} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

All the values in the vector $\vec{x}$ are the same value x.

$$\begin{aligned} b_1 &= \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \end{bmatrix} \begin{bmatrix} x \\ x \\ \vdots \\ x \end{bmatrix} \\ &= a_{11}x + a_{12}x + \cdots + a_{1n}x \\ &= x(a_{11} + a_{12} + \cdots + a_{1n}) \end{aligned}$$

Remember that the values in $\vec{\alpha}_1{}^T$ sum to 1.

$$b_1 = x(a_{11} + a_{12} + \cdots + a_{1n}) = x \times 1 = x$$

As we can see, $b_1 = x$. This property carries through for each row, as all of them sum to 1. Thus $\vec{b} = \vec{x}$, which is what we wanted to prove.

2. Show It

Let n be a positive integer. Let $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ be a set of k linearly dependent vectors in $\mathbb{R}^n$. Show that for any $n \times n$ matrix $\mathbf{A}$, the set $\{\mathbf{A}\vec{v}_1, \mathbf{A}\vec{v}_2, \ldots, \mathbf{A}\vec{v}_k\}$ is a set of linearly dependent vectors. Make sure that you prove this rigorously for all possible matrices $\mathbf{A}$.

Solution: It is given that $\{\vec{v}_i \in \mathbb{R}^n | i = 0, 1, \dots, k\}$ are linearly dependent. This implies that there exist k scalars, $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ from $\mathbb{R}$ that are *not all equal to zero simultaneously* (or equivalently *at least one of them is not equal to zero*), such that

$$\alpha_1 \cdot \vec{v}_1 + \alpha_2 \cdot \vec{v}_2 + \cdots + \alpha_k \cdot \vec{v}_k = \vec{0} \tag{1}$$

By multiplying Equation 1 by $\mathbf{A}$ from the left side, we get

$$\mathbf{A} \cdot (\alpha_1 \cdot \vec{v}_1 + \alpha_2 \cdot \vec{v}_2 + \cdots + \alpha_k \cdot \vec{v}_k) = \mathbf{A} \cdot \vec{0} \tag{2}$$

First, note that $\mathbf{A} \cdot \vec{0} = \vec{0}$. Moreover, if we distribute we get

$$\mathbf{A} \cdot (\alpha_1 \cdot \vec{v}_1) + \mathbf{A} \cdot (\alpha_2 \cdot \vec{v}_2) + \cdots + \mathbf{A} \cdot (\alpha_k \cdot \vec{v}_k) = \vec{0} \tag{3}$$

From associativity of multiplication we get

$$(\mathbf{A} \cdot \alpha_1) \cdot \vec{v}_1 + (\mathbf{A} \cdot \alpha_2) \cdot \vec{v}_2 + \cdots + (\mathbf{A} \cdot \alpha_k) \cdot \vec{v}_k = \vec{0} \tag{4}$$

Since scalar-matrix multiplication is commutative, we get

$$(\alpha_1 \cdot \mathbf{A}) \cdot \vec{v}_1 + (\alpha_2 \cdot \mathbf{A}) \cdot \vec{v}_2 + \cdots + (\alpha_k \cdot \mathbf{A}) \cdot \vec{v}_k = \vec{0} \tag{5}$$

By using associativity of multiplication again, we get

$$\alpha_1 \cdot (\mathbf{A} \cdot \vec{v}_1) + \alpha_2 \cdot (\mathbf{A} \cdot \vec{v}_2) + \cdots + \alpha_k \cdot (\mathbf{A} \cdot \vec{v}_k) = \vec{0} \tag{6}$$

Therefore, the same k scalars, $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ show the linear dependence of the vectors $\{\mathbf{A}\vec{v}_1, \mathbf{A}\vec{v}_2, \dots, \mathbf{A}\vec{v}_k\}$ as requested.

■

Note: There are alternative and equivalent implications of linear dependence that can be used in the proof (in place of Equation 1). Here are a few of them:

(a) The vector $\vec{v}_k$ can be represented as a linear combination of the *other* vectors as follows: There exist scalars $\alpha_1, \alpha_2, \dots, \alpha_{k-1}$ such that

$$\alpha_1 \cdot \vec{v}_1 + \alpha_2 \cdot \vec{v}_2 + \cdots + \alpha_{k-1} \cdot \vec{v}_{k-1} = \vec{v}_k \tag{7}$$

(in this alternative, the scalars may all be zeros, and the linear combination in the left hand side must exclude $\vec{v}_k$)

(b) Some vector $\vec{v}_j$ can be represented as a linear combination of the *other* vectors as follows: There exist scalars $\alpha_i, 1 \leq i \leq k, i \neq j$ such that

$$\sum_{\substack{i=1 \\ i \neq j}}^{k} \alpha_i \cdot \vec{v}_i = \vec{v}_j \tag{8}$$

(in this alternative, the scalars may all be zeros, and the linear combination in the left hand side must exclude $\vec{v}_j$)

(c) There exists $1 \leq j \leq k$ such that $\vec{v}_j$ can be represented as a linear combination of the $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{j-1}$ as follows: There exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_{j-1}$ such that

$$\sum_{i=1}^{j-1} \alpha_i \cdot \vec{v}_i = \vec{v}_j \tag{9}$$

(in this alternative, the scalars may all be zeros, and the linear combination in the left hand side must exclude $\vec{v}_j$)

For any of these alternatives, the rest of the proof is similar to the one demonstrated above (multiply both sides by $\mathbf{A}$ and then apply the linearity of matrix multiplication to push the $\mathbf{A}$ inside the relevant sum and next to the $\vec{v}_i$) and will result in an equation similar to Equation 6 (that matches the alternative chosen).

Common mistakes included:

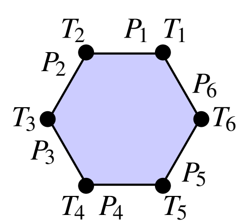
- When using the definition provided in Equation 1, not indicating that the at least one of scalars needs to be different than zero.
- When using the definition provided in Equation 1, stating that all scalars need to be different than zero.
- When using any of the alternative implications (a)-(c) above, requiring that at least one scalar be different than zero (or all different than zero).
- When using any of the alternative implications (a)-(c) above, not excluding the vector on the right hand side from the linear combination on the left hand side.

3. Figuring out the tips

A number of people gather around a round table for a dinner. Between every adjacent pair of people there is a plate for tips. When everyone has finished eating, each person places half their tip in the plate to their left and half in the plate to their right. In the end, of the tips in each plate, some of it is contributed by the person to its right, and the rest is contributed by the person to its left. Suppose you can only see the plates of tips after everyone has left. Can you deduce everyone's individual tip amounts?

Note: For this question, if we assume that tips are positive, we need to introduce additional constraints enforcing that, and we wouldn't get a linear system of equations. So we are going to ignore this constraint and negative tips are ok.

(a) Suppose 6 people sit around a table and there are 6 plates of tips at the end.



If we know the amounts in every plate of tips (P_1 to P_6), can we determine the individual tips of all 6 people (T_1 to T_6)? If yes, explain why. If not, give two different assignments of T_1 to T_6 that will result in the same P_1 to P_6.

Solution:

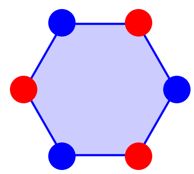
No, this is not possible to determine in general. For example, the following two different assignments of tip amounts for each person:

$$(T_1, T_2, T_3, T_4, T_5, T_6) = (2, 0, 2, 0, 2, 0)$$
$$(T_1, T_2, T_3, T_4, T_5, T_6) = (0, 2, 0, 2, 0, 2)$$

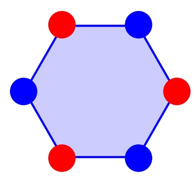
Will both result in $(P_1, P_2, P_3, P_4, P_5, P_6) = (1, 1, 1, 1, 1, 1)$.
If we write down the system of linear equations:

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \\ T_6 \end{bmatrix} = \begin{bmatrix} 2P_1 \\ 2P_2 \\ 2P_3 \\ 2P_4 \\ 2P_5 \\ 2P_6 \end{bmatrix}$$

We can use Gaussian elimination to reduce the last row to all zeros, therefore equations are linearly dependent. However, Gaussian elimination is not needed to solve this question.
Intuitively, we can color each spot on the table alternating between red or blue:

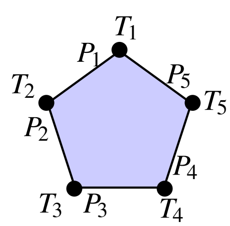


Then supposing that everyone sitting at red spots all tip r dollars and everyone sitting at blue spots all tip b dollars, we find $P_1, \cdots P_6$ dollars on the plates. However, this is no different from this following coloring:



Thus we see that because of the special symmetry of the six-sided table, it's not possible to deduce everyone's tip.

(b) The same question as above, but what if we have 5 people sitting around a table?



Solution: Yes. The problem can be reduced to the following system of equations:

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \end{bmatrix} = \begin{bmatrix} 2P_1 \\ 2P_2 \\ 2P_3 \\ 2P_4 \\ 2P_5 \end{bmatrix}$$

We can then run row-reduction on this matrix. First subtract rows 1 and 3 and add rows 2 and 4 to row 5:

$$\begin{bmatrix}1&1&0&0&0\\0&1&1&0&0\\0&0&1&1&0\\0&0&0&1&1\\0&0&0&0&1\end{bmatrix}\begin{bmatrix}T_1\\T_2\\T_3\\T_4\\T_5\end{bmatrix}=\begin{bmatrix}2P_1\\2P_2\\2P_3\\2P_4\\P_5-P_1+P_2-P_3+P_4\end{bmatrix}$$

Now we subtract row 5 from row 4, row 4 from row 3, row 3 from row 2 and row 2 from row 1:

$$\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix}\begin{bmatrix}T_1\\T_2\\T_3\\T_4\\T_5\end{bmatrix}=\begin{bmatrix}P_1-P_2+P_3-P_4+P_5\\P_2-P_3+P_4-P_5+P_1\\P_3-P_4+P_5-P_1+P_2\\P_4-P_5+P_1-P_2+P_3\\P_5-P_1+P_2-P_3+P_4\end{bmatrix}$$

Thus we have a unique solution for T_1 to T_5 in terms of P_1 to P_5.

Intuitively, unlike the argument for the previous part, since there is an odd number of seats, it's not possible for people to color every alternate seat red or blue.

(c) If n is the total number of people sitting around a table, for which n can you figure out everyone's tip? You do not have to rigorously prove your answer.

Solution: **Note:** Although you didn't need to prove your answers rigorously, here we will give you a rigorous argument. As long as your answer has this flavor of argument (or another equally sound argument), you should give yourself full credit.

For even n, this is not possible. Here is a counterexample: Suppose that all the plates had \$1 in them. There are clearly two different ways that this could have happened. First:

$$T_n=\begin{cases}2 & n \text{ odd}\\0 & n \text{ even}\end{cases}$$

Second:

$$T_n=\begin{cases}0 & n \text{ odd}\\2 & n \text{ even}\end{cases}$$

Which will result in $P_i = 1$ for all i from 1 to n. Therefore we cannot figure out everyone's tip in general.

We can determine everyone's tips for all odd n. You can either argue with Gaussiam elimination on a general $n \times n$ matrix where n is odd, or use the following argument which does not use Gaussian elimination:

Gaussian Elimination Solution:

For odd n, the matrix encoding the system of linear equations is:

$$\begin{array}{cl}\begin{bmatrix}1&1&0&0&\cdots&0\\0&1&1&0&\cdots&0\\&&\ddots&\ddots&&\\0&0&\cdots&1&1&0\\0&0&\cdots&0&1&1\\1&0&\cdots&0&0&1\end{bmatrix} & \begin{array}{l}\text{Row } 1\\\text{Row } 2\\\vdots\\\text{Row } n-2\\\text{Row } n-1\\\text{Row } n\end{array}\end{array}$$

We want to perform Gaussian elimination on this matrix. First we subtract all odd-numbered rows from row n and add all even-numbered rows to row n. What is row n in the end? Denote the i-th item in row n by $R_{n,i}$. We know that $R_{n,1} = 0$, and for $i = 2, \cdots, n-2$, $R_{n,i} = 1 - 1 = 0$. Since n is odd, row $n-2$ is subtracted from row n and row $n-1$ is added to row n. Therefore $R_{n,n-1} = 0$ and $R_{n,n} = 2$. Now we can divide row n by 2, then subtract row n from row $n-1$, row $n-1$ from row $n-2$, and so on, until we get the identity matrix. Therefore we can see that all the rows are linearly independent and we can obtain a unique solution to this system of equations.

Alternate Solution:

Suppose that each customer tipping: $T_1 = a_1, \cdots, T_n = a_n$ gives rise to the amount in plates $P_1 = p_1, \cdots, P_n = p_n$. Now suppose there exist another different solution to $T_1, \cdots, T_n$ that gives the same amount in plates. Then for some i two possible values for T_i, a_i and $a_i + \varepsilon$, are possible. Then:

$$\begin{aligned} a_i + a_{i+1} &= 2p_i \\ T_{i+1} &= 2p_i - T_i \\ &= (a_i + a_{i+1}) - (a_i + \varepsilon) \\ &= a_{i+1} - \varepsilon \end{aligned}$$

This means that $T_{i+1} = a_{i+1} - \varepsilon$, $T_{i+2} = a_{i+2} + \varepsilon$ and so on. We can then keep going around the circle, noting that there are an odd amount of sign flips in total before we get back to T_i. Therefore eventually we get $T_i = a_i - \varepsilon$. However we are assuming that $T_i = a_i + \varepsilon$ in the beginning. Therefore ε must be zero and we conclude that there is one unique solution.

4. Image Stitching

Often, when people take pictures of a large object, they are constrained by the field of view of the camera. This means that they have two options by which they can capture the entire object:

- Stand as far as away as they need to to include the entire object in the camera's field of view (clearly, we do not want to do this as it reduces the amount of detail in the image)
- (This is more exciting) Take several pictures of different parts of the object, and stitch them together, like a jigsaw puzzle.

We are going to explore the second option in this problem. Prof. Ayazifar, who is a professional photographer, wants to construct an image by doing this "image stitching". Unfortunately, he took some of the pictures at different angles, as well as at different positions and distances from the object. While processing these pictures, he lost information about the positions and orientations at which he took them. Luckily, you and your friend Marcela, with your wealth of newly acquired knowledge about vectors and rotation matrices, can help him!

You and Marcela are designing an iPhone app that stitches photographs together into one larger image. Marcela has already written an algorithm that finds common points in overlapping images and it's your job to figure out how to stitch the images together. You recently learned about vectors and rotation matrices in EE16A and you have an idea about how to do this.

Your idea is that you should be able to find a single rotation matrix, $\mathbf{R}$, which is a function of some angle, θ, and a translation vector, $\vec{T}$, that transforms every common point in one image to that same point in the other image. Once you find the the angle, θ, and the translation vector, $\vec{T}$, you will be able to transform one image so that it lines up with the other image.

Suppose $\vec{p}$ is a point in one image and $\vec{q}$ is the corresponding point (i.e. they represent the same thing in the scene) in the other image. You write down the following relationship between $\vec{p}$ and $\vec{q}$.

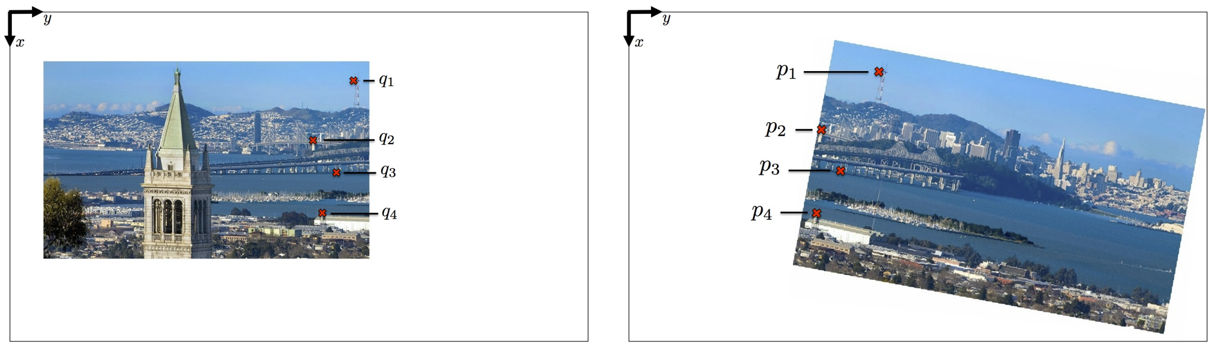


Figure 1: Two images to be stitched together with pairs of matching points labeled.

$$\begin{bmatrix} q_x \\ q_y \end{bmatrix} = \underbrace{\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}}_{\mathbf{R}(\theta)} \begin{bmatrix} p_x \\ p_y \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \end{bmatrix} \tag{10}$$

This looks good but then you realize that one of the pictures might be farther away than the other. You realize that you need to add a scaling factor, $\lambda > 0$.

$$\begin{bmatrix} q_x \\ q_y \end{bmatrix} = \lambda \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} p_x \\ p_y \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \end{bmatrix} \tag{11}$$

(For example, if $\lambda > 1$, then the image containing q is closer (appears larger) than the image containing p. If $0 < \lambda < 1$, then the image containing q appears smaller.)

You are now confident that if you can find θ, $\vec{T}$, and λ, then you will be able to reorient and scale one of the images so that it lines up with the other image.

Before you get too excited, however, you realize that you have a problem. Equation (11) is not a linear equation in θ, $\vec{T}$, and λ. You're worried that you don't have a good technique for solving nonlinear systems of equations. You decide to talk to Marcela and the two of you come up with a brilliant solution.

You decide to "relax" the problem so that you're solving for a general matrix $\mathbf{R}$ rather than precisely a scaled rotation matrix. The new equation you come up with is

$$\begin{bmatrix} q_x \\ q_y \end{bmatrix} = \begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix} \begin{bmatrix} p_x \\ p_y \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \end{bmatrix} \tag{12}$$

This equation is linear so you can solve for R_{xx}, R_{xy}, R_{yx}, R_{yy}, T_x, T_y. Also you realize that if $\vec{p}$ and $\vec{q}$ actually do differ by a rotation of θ degrees and a scaling of λ, you can expect that the general matrix $\mathbf{R}$ that you find will turn out to be a scaled rotation matrix with $R_{xx} = \lambda\cos(\theta)$, $R_{xy} = -\lambda\sin(\theta)$, $R_{yx} = \lambda\sin(\theta)$, and $R_{yy} = \lambda\cos(\theta)$.

(a) Multiply out Equation (12) into two scalar linear equations. What are the known values and what are the unknowns in each equation? How many unknowns are there? How many equations do you need to solve for all the unknowns? How many pairs of common points $\vec{p}$ and $\vec{q}$ will you need in order to write down a system of equations that you can use to solve for the unknowns?

Solution: We can rewrite the above matrix equation as the two scalar linear equations

$$q_x = p_x R_{xx} + p_y R_{xy} + T_x \tag{13}$$
$$q_y = p_x R_{yx} + p_y R_{yy} + T_y \tag{14}$$

Here the known values are the elements of the pair of points: $q_x,\ q_y,\ p_x,\ p_y$, and 1. The unknowns are elements of $\mathbf{R}$ and $\vec{T}$: $R_{xx},\ R_{xy},\ R_{yx},\ R_{yy},\ T_x$, and T_y. There are 6 unknowns so we need a total of 6 equations to solve for them. For every pair of points we add we get two more equations. Thus, we need 3 pairs of common points to get 6 equations.

(b) Write out a system of linear equations that you can use to solve for the values of $\mathbf{R}$ and $\vec{T}$.

Solution: We will label the 3 pairs of point we select as

$$\vec{q}_1 = \begin{bmatrix} q_{1x} \\ q_{1y} \end{bmatrix}, \quad \vec{p}_1 = \begin{bmatrix} p_{1x} \\ p_{1y} \end{bmatrix} \qquad \vec{q}_2 = \begin{bmatrix} q_{2x} \\ q_{2y} \end{bmatrix}, \quad \vec{p}_2 = \begin{bmatrix} p_{2x} \\ p_{2y} \end{bmatrix} \qquad \vec{q}_3 = \begin{bmatrix} q_{3x} \\ q_{3y} \end{bmatrix}, \quad \vec{p}_3 = \begin{bmatrix} p_{3x} \\ p_{3y} \end{bmatrix} \tag{15}$$

We write out the system of linear equations in matrix form.

$$\begin{bmatrix} p_{1x} & p_{1y} & 0 & 0 & 1 & 0 \\ 0 & 0 & p_{1x} & p_{1y} & 0 & 1 \\ p_{2x} & p_{2y} & 0 & 0 & 1 & 0 \\ 0 & 0 & p_{2x} & p_{2y} & 0 & 1 \\ p_{3x} & p_{3y} & 0 & 0 & 1 & 0 \\ 0 & 0 & p_{3x} & p_{3y} & 0 & 1 \end{bmatrix} \begin{bmatrix} R_{xx} \\ R_{xy} \\ R_{yx} \\ R_{yy} \\ T_x \\ T_y \end{bmatrix} = \begin{bmatrix} q_{1x} \\ q_{1y} \\ q_{2x} \\ q_{2y} \\ q_{3x} \\ q_{3y} \end{bmatrix} \tag{16}$$

(c) In the IPython notebook `prob2.ipynb` you will have a chance to test out your solution. Plug in the values that you are given for $p_x,\ p_y,\ q_x$, and q_y for each pair of points into your system of equations to solve for the parameters $\mathbf{R}$ and $\vec{T}$. You will be prompted to enter your results and the notebook will then apply your transformation to the second image and show you if your stitching algorithm works.

Solution: The parameters for the transformation from the coordinates of the first image to those of the second image are $\mathbf{R} = \begin{bmatrix} 1.1954 & .1046 \\ -.1046 & 1.1954 \end{bmatrix}$ (this corresponds to $\lambda = 1.2$, $\theta = -5$ degrees), and $\vec{T} = \begin{bmatrix} -150 \\ -250 \end{bmatrix}$.

(d) We will now explore when this algorithm fails. For example, the three pairs of points must all be distinct points. Show that if $\vec{p}_1, \vec{p}_2, \vec{p}_3$ are *co-linear*, the system of (12) is underdetermined. Does this make sense geometrically?

(Think about the kinds of transformations possible by a general affine transform. An affine transform is one that preserves points, for example in the rotation of a line although the angle of the line might change the length will not. All linear transformations are affine. **Click me! Definition of Affine.**)

Use the fact that: $\vec{p}_1, \vec{p}_2, \vec{p}_3$ are co-linear iff $(\vec{p}_2 - \vec{p}_1) = k(\vec{p}_3 - \vec{p}_1)$ for some $k \in \mathbb{R}$.

Solution:

*Note: A general 2D affine transformation is a transformation from points $\vec{p}$ to points $\vec{q}$ of the form of Equation (12). It is a linear transform (by matrix **R**) followed by a translation by vector $\vec{T}$.*

The algorithm fails when the points p_1, p_2, p_3 are not all distinct, or when they are co-linear.

To show that the co-linear case leads to an underdetermined system, write the system (in matrix variable R) as:

$$\begin{cases} \vec{q}_1 = \mathbf{R}\vec{p}_1 + \vec{T} \\ \vec{q}_2 = \mathbf{R}\vec{p}_2 + \vec{T} \\ \vec{q}_3 = \mathbf{R}\vec{p}_3 + \vec{T} \end{cases}$$

By subtracting the first equation from the last two, this is equivalent to the following system:

$$\begin{cases} \vec{q}_1 = \mathbf{R}\vec{p}_1 + \vec{T} \\ \vec{q}_2 - \vec{q}_1 = \mathbf{R}(\vec{p}_2 - \vec{p}_1) \\ \vec{q}_3 - \vec{q}_1 = \mathbf{R}(\vec{p}_3 - \vec{p}_1) \end{cases}$$

Then, by co-linearity, this is equivalent to the system:

$$\begin{cases} \vec{q}_1 = \mathbf{R}\vec{p}_1 + \vec{T} \\ \vec{q}_2 - \vec{q}_1 = \mathbf{R}k(\vec{p}_3 - \vec{p}_1) \\ \vec{q}_3 - \vec{q}_1 = \mathbf{R}(\vec{p}_3 - \vec{p}_1) \end{cases}$$

Notice that the last two equations are now linearly dependent. To see this, remember that the variables are the entries of the matrix $\mathbf{R} = \begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix}$, and all other elements $(\vec{p}_i, \vec{q}_i, k)$ are constants. Define $\vec{p}_{31} = \vec{p}_3 - \vec{p}_1$. The third equation is a constraint on linear combinations of entries in $\mathbf{R}$ – specifically, it is a constraint on $\mathbf{R}\vec{p}_{31}$. The second equation is a constraint on $\mathbf{R}(k\vec{p}_{31})$. However, since $\mathbf{R}(k\vec{p}_{31}) = k(\mathbf{R}\vec{p}_{31})$, these two constraints are linearly dependent. Therefore the system is underdetermined.

This problem can also be solved by writing out the linear equations explicitly in matrix form, as you did for part (b). The linear system is:

$$\underbrace{\begin{bmatrix} p_{1x} & p_{1y} & 0 & 0 & 1 & 0 \\ 0 & 0 & p_{1x} & p_{1y} & 0 & 1 \\ p_{2x} & p_{2y} & 0 & 0 & 1 & 0 \\ 0 & 0 & p_{2x} & p_{2y} & 0 & 1 \\ p_{3x} & p_{3y} & 0 & 0 & 1 & 0 \\ 0 & 0 & p_{3x} & p_{3y} & 0 & 1 \end{bmatrix}}_{A} \begin{bmatrix} R_{xx} \\ R_{xy} \\ R_{yx} \\ R_{yy} \\ T_x \\ T_y \end{bmatrix} = \begin{bmatrix} q_{1x} \\ q_{1y} \\ q_{2x} \\ q_{2y} \\ q_{3x} \\ q_{3y} \end{bmatrix}$$

Subtracting row 1 from rows 4 and 5, then subtracting row 2 from rows 4 and 6, this is equivalent to the system specified by matrix:

$$\begin{bmatrix} p_{1x} & p_{1y} & 0 & 0 & 1 & 0 \\ 0 & 0 & p_{1x} & p_{1y} & 0 & 1 \\ p_{2x} - p_{1x} & p_{2y} - p_{1y} & 0 & 0 & 0 & 0 \\ 0 & 0 & p_{2x} - p_{1x} & p_{2y} - p_{1y} & 0 & 0 \\ p_{3x} - p_{1x} & p_{3y} - p_{1y} & 0 & 0 & 0 & 0 \\ 0 & 0 & p_{3x} - p_{1x} & p_{3y} - p_{1y} & 0 & 0 \end{bmatrix}$$

Then, by co-linearity, we have $p_{2x} - p_{1x} = k(p_{3x} - p_{1x})$ and so on, so the above matrix is:

$$\begin{bmatrix} p_{1x} & p_{1y} & 0 & 0 & 1 & 0 \\ 0 & 0 & p_{1x} & p_{1y} & 0 & 1 \\ k(p_{3x} - p_{1x}) & k(p_{3y} - p_{1y}) & 0 & 0 & 0 & 0 \\ 0 & 0 & k(p_{3x} - p_{1x}) & k(p_{3y} - p_{1y}) & 0 & 0 \\ p_{3x} - p_{1x} & p_{3y} - p_{1y} & 0 & 0 & 0 & 0 \\ 0 & 0 & p_{3x} - p_{1x} & p_{3y} - p_{1y} & 0 & 0 \end{bmatrix}$$

Now clearly rows 3 and 5 are linearly dependent (and in fact, so are rows 4 and 6). Therefore the linear system is underdetermined.

Geometrically: Consider a transform that simply stretches along the x-axis. For example, $\vec{T} = 0$ and $\mathbf{R} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$. Say we pick all $\vec{p}_i$ to lie along the y-axis. Then applying this transformation will send $\vec{q}_i = \vec{p}_i$ (since the y-axis is unaffected by this stretching). However, only given q_i, this transform has the same effect as the identity transform (that is, where $\vec{T} = 0$ and $\mathbf{R}$ is the identity). Therefore the transform cannot be determined from the q_i (both the identity transform and the x-axis stretching lead to the same q_i). This argument can be extended to the case when p_i are co-linear, but not necessarily along the x-axis.

(e) *(PRACTICE)* Show that if the three points are not co-linear, the system is fully determined.

Solution: As before, write the system (in matrix variable $\mathbf{R}$) as:

$$\begin{cases} \vec{q}_1 = \mathbf{R}\vec{p}_1 + \vec{T} \\ \vec{q}_2 = \mathbf{R}\vec{p}_2 + \vec{T} \\ \vec{q}_3 = \mathbf{R}\vec{p}_3 + \vec{T} \end{cases}$$

We can subtract the first equation from the last two, to find:

$$\begin{cases} \vec{q}_2 - \vec{q}_1 = \mathbf{R}(\vec{p}_2 - \vec{p}_1) \\ \vec{q}_3 - \vec{q}_1 = \mathbf{R}(\vec{p}_3 - \vec{p}_1) \end{cases}$$

Define $\vec{p}_{21} = \vec{p}_2 - \vec{p}_1$ and $\vec{p}_{31} = \vec{p}_3 - \vec{p}_1$, and similarly $\vec{q}_{21} = \vec{q}_2 - \vec{q}_1$ and $\vec{q}_{31} = \vec{q}_3 - \vec{q}_1$. Then the above system can be written as:

$$\begin{bmatrix} \vec{q}_{21} & \vec{q}_{31} \end{bmatrix} = R \begin{bmatrix} \vec{p}_{21} & \vec{p}_{31} \end{bmatrix} \tag{17}$$

Or, writing $\mathbf{R}$ explicitly:

$$\begin{bmatrix} \vec{q}_{21} & \vec{q}_{31} \end{bmatrix} = \begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix} \begin{bmatrix} \vec{p}_{21} & \vec{p}_{31} \end{bmatrix} \tag{18}$$

This can be written as a system of two equations, one for each row of the matrix $\mathbf{R}$. For example,

$$\vec{\alpha}_1^T = \begin{bmatrix} R_{xx} & R_{xy} \end{bmatrix} \begin{bmatrix} \vec{p}_{21} & \vec{p}_{31} \end{bmatrix}$$

Where $\vec{\alpha}_1^T$ is the first row of matrix the $\begin{bmatrix} \vec{q}_{21} & \vec{q}_{31} \end{bmatrix}$. Then, transposing:

$$\begin{bmatrix} \vec{p}_{21} & \vec{p}_{31} \end{bmatrix}^T \begin{bmatrix} R_{xx} \\ R_{xy} \end{bmatrix} = \vec{\alpha}_1$$

Notice that by non-co-linearity, the vectors $\vec{p}_{21}$ and $\vec{p}_{31}$ are not linearly dependent. Therefore the (2×2) matrix $\begin{bmatrix} \vec{p}_{21} & \vec{p}_{31} \end{bmatrix}^T$ has linearly independent rows, so this system can be solved for R_{xx} and R_{xy}, the first row of $\mathbf{R}$. The second row of $\mathbf{R}$ can be found analogously.

Then we can find $\vec{T}$, by using one of the correspondence pairs: $\vec{T} = \vec{q}_1 - \mathbf{R}\vec{p}_1$

(f) *(PRACTICE)* Marcela comments that perhaps the system (with three co-linear points) is only underdetermined because we “relaxed” our model too much by allowing for general affine transforms, instead of just isotropic-scale/rotation/translation. Can you come up with a different representation of (11), that will allow for recovering the transform from only *two* pairs of distinct points?
(Hint: Let $a = \lambda\cos(\theta)$ and $b = \lambda\sin(\theta)$. In other words, enforce $R_{xx} = R_{yy}$ and $R_{xy} = -R_{yx}$).

Solution: Model the system as:

$$\vec{q} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \vec{p} + \vec{T}$$

in variables $a, b \in \mathbb{R}$. Given two pairs of points p_i, q_i, we have the system:

$$\begin{cases} \vec{q}_1 = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \vec{p}_1 + \vec{T} \\ \vec{q}_2 = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \vec{p}_2 + \vec{T} \end{cases}$$

We can subtract to find:

$$\vec{q}_2 - \vec{q}_1 = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} (\vec{p}_2 - \vec{p}_1)$$

Define $\vec{v} = \vec{q}_2 - \vec{q}_1$ and $\vec{w} = \vec{p}_2 - \vec{p}_1$. Then, writing the above system in components:

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} w_x \\ w_y \end{bmatrix}$$

This can be written explicitly as a linear system in variables a, b as:

$$\underbrace{\begin{bmatrix} w_x & -w_y \\ w_y & w_x \end{bmatrix}}_{W} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} v_x \\ v_y \end{bmatrix}$$

If $\vec{p}_1 \neq \vec{p}_2$, then $\vec{w} \neq \vec{0}$, and the matrix W has linearly independent rows and columns. Therefore we can solve for a, b given only two pairs of distinct points.

5. Homework process and study group

Who else did you work with on this homework? List names and student ID’s. (In case of hw party, you can also just describe the group.) How did you work on this homework?
Working in groups of 3-5 will earn credit for your participation grade.

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.

6. (PRACTICE) Powers of a Nilpotent Matrix

Do this problem if you would like more mechanical practice with matrix multiplication.
The following matrices are examples of a special type of matrix called a nilpotent matrix. What happens to each of these matrices when you multiply it by itself four times? Multiply them to find out. Why do you think these are called "nilpotent" matrices? (Of course, there is nothing magical about 4×4 matrices. You can have nilpotent square matrices of any dimension greater than 1.)

(a) Do $\mathbf{A}^4$ by hand. Make sure you show what $\mathbf{A}^2$ and $\mathbf{A}^3$ are along the way.

$$\mathbf{A} = \begin{bmatrix} 0 & 2 & 1 & 3 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{19}$$

Solution:

$$\mathbf{A}^2 = \begin{bmatrix} 0 & 2 & 1 & 3 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 & 3 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{20}$$

$$\mathbf{A}^3 = \begin{bmatrix} 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 & 3 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{21}$$

$$\mathbf{A}^4 = \begin{bmatrix} 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 & 3 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{22}$$

(b) Do $\mathbf{B}^4$ by hand. Make sure you show what $\mathbf{B}^2$ and $\mathbf{B}^3$ are along the way.

$$\mathbf{B} = \begin{bmatrix} 3 & 4 & 2 & 1 \\ -5 & -6 & -3 & -1 \\ 6 & 7 & 3 & 2 \\ 2 & 2 & 1 & 0 \end{bmatrix} \tag{23}$$

Solution:

$$\mathbf{B}^2 = \begin{bmatrix} 3 & 4 & 2 & 1 \\ -5 & -6 & -3 & -1 \\ 6 & 7 & 3 & 2 \\ 2 & 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 4 & 2 & 1 \\ -5 & -6 & -3 & -1 \\ 6 & 7 & 3 & 2 \\ 2 & 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 4 & 1 & 3 \\ -5 & -7 & -2 & -5 \\ 5 & 7 & 2 & 5 \\ 2 & 3 & 1 & 2 \end{bmatrix} \tag{24}$$

$$\mathbf{B}^3 = \begin{bmatrix} 3 & 4 & 1 & 3 \\ -5 & -7 & -2 & -5 \\ 5 & 7 & 2 & 5 \\ 2 & 3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 3 & 4 & 2 & 1 \\ -5 & -6 & -3 & -1 \\ 6 & 7 & 3 & 2 \\ 2 & 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ -2 & -2 & 0 & -2 \\ 2 & 2 & 0 & 2 \\ 1 & 1 & 0 & 1 \end{bmatrix} \tag{25}$$

$$\mathbf{B}^4 = \begin{bmatrix} 1 & 1 & 0 & 1 \\ -2 & -2 & 0 & -2 \\ 2 & 2 & 0 & 2 \\ 1 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 4 & 2 & 1 \\ -5 & -6 & -3 & -1 \\ 6 & 7 & 3 & 2 \\ 2 & 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{26}$$

A nilpotent matrix is a matrix that becomes all 0's when you raise it to some power, i.e. repeatedly multiply it by itself.

7. **(PRACTICE) Elementary Matrices**

Last week, we learned about an important technique for solving systems of linear equations called Gaussian Elimination. It turns out each row operation in Gaussian Elimination can be performed by multiplying the augmented matrix on the left by a specific matrix called an *elementary matrix*. For example, suppose we want to row reduce the following augmented matrix:

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 0 & -5 & \vdots & 16 \\ 0 & 1 & 0 & 3 & \vdots & -7 \\ -2 & -3 & 1 & -6 & \vdots & 9 \\ 0 & 1 & 0 & 2 & \vdots & -5 \end{bmatrix} \tag{27}$$

What matrix do you get when you subtract the 4th row from the 2nd row of $\mathbf{A}$ (putting the result in row 2)? (You don't have to include this in your solutions.) Now, try multiplying the original $\mathbf{A}$ on the left by

$$\mathbf{E} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{28}$$

(You don't have to include this in your solutions either.) Notice that you get the same thing.

$$\mathbf{EA} = \begin{bmatrix} 1 & -2 & 0 & -5 & \vdots & 16 \\ 0 & 0 & 0 & 1 & \vdots & -2 \\ -2 & -3 & 1 & -6 & \vdots & 9 \\ 0 & 1 & 0 & 2 & \vdots & -5 \end{bmatrix} \tag{29}$$

$\mathbf{E}$ is a special type of matrix called an *elementary matrix*. This means that we can obtain the matrix $\mathbf{E}$ from the identity matrix by applying an elementary row operation - in this case, subtracting the 4th row from the 2nd row.

In general, any elementary row operation can be performed by left multiplying by an appropriate elementary matrix. In other words, you can perform a row operation on a matrix $\mathbf{A}$ by first performing that row operation on the identity matrix to get an elementary matrix, and then left multiplying $\mathbf{A}$ by the elementary matrix (like we did above).

(a) Write down the elementary matrices required to perform the following row operations on a 4×5 augmented matrix.

- Switching rows 1 and 2
- Multiplying row 3 by -4
- Adding $2\times$ row 2 to row 4 (putting the result in row 4) and subtracting row 2 from row 1 (putting the result in row 1)

Hint: For this last problem, note that if you want to perform two row operations on the matrix $\mathbf{A}$, you can perform them both on the identity matrix and then left multiply $\mathbf{A}$ by the resulting matrix.

Solution: We obtain each of the desired elementary matrices by performing the row operations on a 4×4 identity matrix

- Switching rows 1 and 2:

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{30}$$

- Multiplying row 3 by -4:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{31}$$

- Adding $2\times$ row 2 to row 4 and subtracting row 2 from row 1:

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{bmatrix} \tag{32}$$

  Note that we obtained this last matrix by applying two elementary row operations to the identity matrix. We could have performed each elementary row operation on individual identity matrices and then multiplied them together to achieve the same result. In this problem, the order of the matrices did not matter; however, this is not true in general.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{bmatrix} \tag{33}$$

(b) Now, compute a matrix $\mathbf{E}$ (by hand) that fully row reduces the augmented matrix $\mathbf{A}$ given in Eqn (27) - that is find $\mathbf{E}$ such that $\mathbf{EA}$ is in full row reduced echelon form. Show that this is true by multiplying out $\mathbf{EA}$. As a reminder in this case, when the augmented matrix is fully row-reduced it will have the form

$$\mathbf{EA} = \begin{bmatrix} 1 & 0 & 0 & 0 & \vdots & b_1 \\ 0 & 1 & 0 & 0 & \vdots & b_2 \\ 0 & 0 & 1 & 0 & \vdots & b_3 \\ 0 & 0 & 0 & 1 & \vdots & b_4 \end{bmatrix} \tag{34}$$

Hint: As before note that you can either ***apply a set of row operations to the same identity matrix*** *or* ***apply them to separate identity matrices and then multiply the matrices together****. Make sure, though, that you both apply the row operations in the correct order and multiply the matrices in the correct order.*

Solution: We first need to row reduce $\mathbf{A}$ by hand to find the set of required row operations. The following row operations will do the trick (though you could have used a different set that does the same thing.)

- Step 1: Add $2\times$ **Row 1** to Row 3

- Step 2: Add $2\times$ **Row 2** to Row 1, add $7\times$ **Row 2** to Row 3, and subtract **Row 2** from Row 4
- Step 3: Add **Row 4** to Row 1, add $3\times$ **Row 4** to Row 2, and add $5\times$ **Row 4** to Row 3.
- Step 4: Multiply Row 4 by -1

Note that we have grouped row operations together so that each step involves adding a scalar multiple of a particular row to the other rows. This will make calculating the elementary matrices for each step particularly easy.

Applying each of these sets of row operations to a 4×4 identity matrix gives us the following matrices

- Step 1:

$$\mathbf{E}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{35}$$

- Step 2:

$$\mathbf{E}_2 = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 7 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix} \tag{36}$$

- Step 3:

$$\mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{37}$$

- Step 4:

$$\mathbf{E}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{38}$$

We now multiply these matrices together as follows

$$\mathbf{E} = \mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1 = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 1 \end{bmatrix}}\underbrace{\begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 7 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}} \tag{39}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 7 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix} \tag{40}$$

$$= \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & -2 & 0 & 3 \\ 2 & 2 & 1 & 5 \\ 0 & 1 & 0 & -1 \end{bmatrix} \tag{41}$$

Note the order in which we multiplied the matrices. $\mathbf{E}_1$ gets applied first so it is furthest to the right (i.e. it will act on the augmented matrix first), etc. Also, note that we could have applied the row operations to the identity matrices in different groups. For example, we could have written an elementary matrix for each individual row operation and multiplied all of them together, making sure to maintain the correct order. We also could have applied all of the row operations, in the correct order, to a single identity matrix. **The important thing is that we maintain the correct order of row operations - either when we're applying them to an individual identity matrix or multiplying elementary matrices together.**

To show that $\mathbf{E}$ does, in fact, row reduce $\mathbf{A}$, we calculate $\mathbf{EA}$.

$$\mathbf{EA} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & -2 & 0 & 3 \\ 2 & 2 & 1 & 5 \\ 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 0 & -5 & \vdots & 16 \\ 0 & 1 & 0 & 3 & \vdots & -7 \\ -2 & -3 & 1 & -6 & \vdots & 9 \\ 0 & 1 & 0 & 2 & \vdots & -5 \end{bmatrix} \tag{42}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 & \vdots & 4 \\ 0 & 1 & 0 & 0 & \vdots & -1 \\ 0 & 0 & 1 & 0 & \vdots & 2 \\ 0 & 0 & 0 & 1 & \vdots & -2 \end{bmatrix} \tag{43}$$