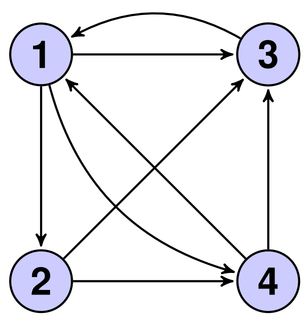
EECS 16A Designing Information Devices and Systems I
Spring 2023 Lecture Notes Note 9

9.1 Page Rank: Eigenvalues and Eigenvectors in Action

Google's Page Rank algorithm uses a web crawler to estimate the popularity of webpages. We assume that the more popular a website, the more incoming links it has. We use the following setup to model the problem: suppose we have a large population of web surfers scattered at random around the internet. At each time step, all surfers randomly choose an outgoing link on their current page to arrive at a new webpage. After a large number of time steps, we expect the number of viewers on each page to reflect that page's popularity.

It's important to realize that this is a simplified model that leaves out certain characteristics of the real world: for example, people do not click links at random and people often seek out entirely new pages rather than clicking on outgoing links. Page Rank as it's used in practice, while still a simplified model, has many more intricacies than we cover here. However, even this first order approximation is a very powerful tool.

Let's look at an internet consisting of four webpages:



How can we calculate the number of viewers on each page after infinite time steps? In this note, we will introduce three important new concepts to help us do this calculation: **Determinants**, **Eigenvalues**, and **Eigenvectors**.

Let's say we started off with 6 people on each webpage. Then one time step ticks and everyone randomly selects an outgoing link to visit. How many people do we expect to be on each webpage now?

Page 1: Page 3 only links to page 1, so all 6 people currently on 3 will move to 1. Page 4 links to pages 1 and 3, so on average about half the people on page 4 will go to page 1. There are no other incoming links, so we expect page 1 to have about $6 \times 1 + 6 \times \frac{1}{2} = 9$ people in the next time step.

Page 2: Page 1 is the only page that links to page 2. We expect a third of people from page 1 to move to page 2, so we expect page 2 to have $6 \times \frac{1}{3} = 2$ people in the next time step.

Page 3: We can arrive at page 3 from all the other pages. At page 1, the fraction of the people who move to page 3 is $\frac{1}{3}$, at page 2 the fraction of the people who move to page 3 is $\frac{1}{2}$, and at page 4 the fraction of the people who move to page 3 is $\frac{1}{2}$. Hence, we expect $6\times\frac{1}{3}+6\times\frac{1}{2}+6\times\frac{1}{2}=8$ people to be on page 3 in the next time step.

Page 4: We can arrive at page 4 from either page 1 or page 2. At page 1, the fraction of people who move to page 4 is $\frac{1}{3}$; at page 2, the fraction of people who move to page 4 is $\frac{1}{2}$. We expect $6\times\frac{1}{3}+6\times\frac{1}{2}=5$ people to be on page 4 in the next time step.

Notice that to find the number of people we expect to be on page i, we added up the number of people we expected would move from every other page to page i. In general, if we have n pages, $x_i(k)$ is the number of people on page i at time step k, and $p(x,y)$ is the fraction of people that will jump from page x to page y, then

$$x_i(k+1) \;=\; \sum_{j=1}^{n} x_j(k)\, p(j,i)$$

If we want to compute the expected number of viewers at time $k+1$ for all pages simultaneously, we can use matrix notation as we did with the pump and reservoir system we looked at earlier. Define $\vec{x}(k) = \begin{bmatrix} x_1(k) & x_2(k) & \dots & x_n(k)\end{bmatrix}^T$ as the vector encoding the number of viewers on each page at time k. Then

$$\vec{x}(k+1) \;=\; \begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ \vdots \\ x_n(k+1)\end{bmatrix} \;=\; \begin{bmatrix} p(1,1) & p(2,1) & \dots & p(n,1) \\ p(1,2) & p(2,2) & \dots & p(n,2) \\ \vdots & \vdots & \vdots & \vdots \\ p(1,n) & p(2,n) & \dots & p(n,n)\end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_n(k)\end{bmatrix}$$

We call the matrix of fractions P. Check that the entry $x_i(k+1)$ does in fact match the summation above. For our specific example, the matrix equation looks like this:

$$\begin{bmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{bmatrix} \begin{bmatrix} 6 \\ 6 \\ 6 \\ 6 \end{bmatrix} \;=\; \begin{bmatrix} 9 \\ 2 \\ 8 \\ 5 \end{bmatrix}$$

If we start with initial counts $\vec{x}(0)$ and want to find the expected number of viewers on each page at time k, we compute

$$\vec{x}(k) \;=\; P^k\,\vec{x}(0) \;=\; \begin{bmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{bmatrix}^k \vec{x}(0)$$

Since we only care about the relative popularity of each webpage, we can think of $\vec{x}(k)$ as fractions of viewers on each page (instead of the total number of viewers). We assume the viewers are initially distributed equally on all webpages, and we initialize $\vec{x}(0)$ as $\begin{bmatrix}\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4}\end{bmatrix}^T$. After running the simulation for many time steps, we hope $\vec{x}(k)$ will reflect the popularity of the webpages. This system is only useful to us if the values of $\vec{x}(k)$ *converge* to some stable fractions, i.e. the fraction of people on each webpage does not change at every timestep.

Once our fractions exactly match these "stable fractions", running the simulation for another time step should not change the values. This means that if the fractions converge, at some point we'll reach fractions $\vec{x}^*$ such that

$$\vec{x}^* = P\vec{x}^*$$

9.2 Eigenvectors and Eigenvalues

In our Page Rank example, $\vec{x}^*$ is an example of an *eigenvector* of P. But eigenvectors have a more general definition:

Definition 9.1 (Eigenvectors and Eigenvalues): Consider a square matrix $A \in \mathbb{R}^{n\times n}$. An **eigenvector** of A is a *nonzero* vector $\vec{x} \in \mathbb{R}^n$ such that

$$A\vec{x} = \lambda\vec{x}$$

where λ is a scalar value, called the **eigenvalue** of $\vec{x}$.

How do we solve this equation for both $\vec{x}$ and λ, knowing only A? It seems we don't have enough information. We start by bringing everything over to one side:

$$A\vec{x} - \lambda\vec{x} = \vec{0}$$

We would like to factor out $\vec{x}$ on the left side, but currently the dimensions don't agree. $\vec{x}$ is an $n\times 1$ vector, while A is an $n\times n$ matrix. To fix this, we replace $\lambda\vec{x}$ with $\lambda I_n\vec{x}$, where I_n is the $n\times n$ identity matrix:

$$(A - \lambda I_n)\vec{x} = \vec{0} \tag{1}$$

Remember, in the definition of an eigenvector, we explicitly excluded the $\vec{0}$ vector. We know at least one element of $\vec{x}$ is nonzero, yet $(A - \lambda I_n)\vec{x} = \vec{0}$. To see what this means, let's define the columns of $A - \lambda I_n$ to be the vectors $v_1 \dots v_n$ and rewrite the product in terms of the columns:

$$\begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} | \\ \vec{v}_1 \\ | \end{bmatrix} x_1 + \begin{bmatrix} | \\ \vec{v}_2 \\ | \end{bmatrix} x_2 + \dots + \begin{bmatrix} | \\ \vec{v}_n \\ | \end{bmatrix} x_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

From this formulation, we can see that some nonzero linear combination of the columns of $A - \lambda I_n$ results in $\vec{0}$. This means that the columns of this matrix must be linearly dependent. Now we can rephrase our problem: For what values of λwill $A - \lambda I_n$ have linearly dependent columns? To help us answer this question, we introduce the determinant.

> **Additional Resources** For more on eigenvalues and eigenvectors, read *Strang* pages 288 - 291 or read *Schuam's* pages 296-299.

9.3 Determinants

Every square matrix has a quantity associated with it known as the *determinant*. This quantity encodes many important properties of the matrix, and it is also intimately connected to its eigenvalues.

For this class, we only need to know how to compute the determinant of a 2×2 matrix by hand. This is given by

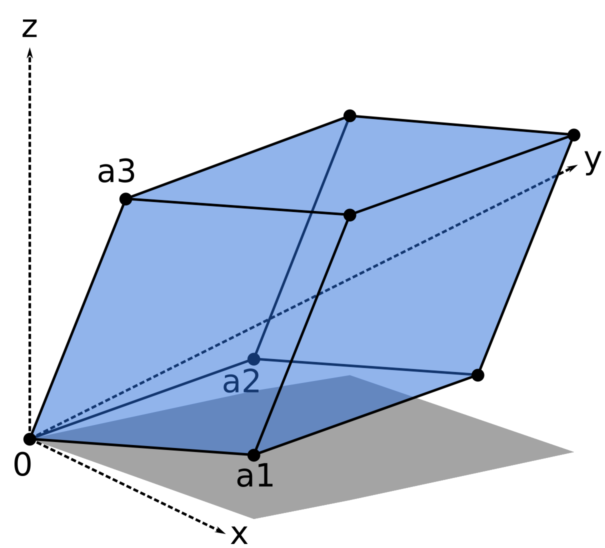
$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc$$

.

There is a beautiful recursive definition for determinants of general $n\times n$ matrices which is outside the scope of this class, but you can read the Wikipedia article on determinants if you want to learn more about it.

Suppose we have a $n\times n$ square matrix $A-\lambda I_n$ and a non-zero vector $x\in\mathbb{R}^n$. Recall from the previous section that the columns of $A-\lambda I_n$ are linearly dependent. We will show that if a matrix's columns are linearly dependent, then its determinant is zero.

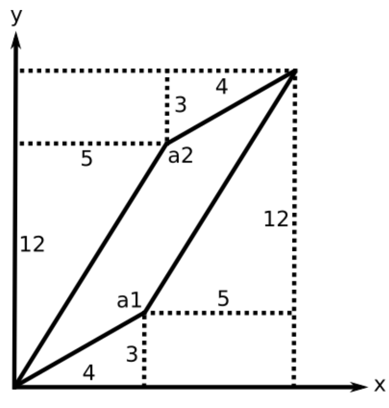
Let's consider the geometric connection between the column vectors of a matrix and its determinant. For simplicity, first consider a 3×3 matrix. Let its column vectors be a_1, a_2 and a_3. Then consider the parallelepiped formed by these vectors: [1]



We define the determinant of this matrix as the volume of this parallelepiped formed by its column vectors. For a 2×2 matrix, its determinant is the area of the parallelogram generated by its 2 column vectors.

[1] Modified from image by Claudio Rocchini, `https://en.wikipedia.org/wiki/Determinant`

Similarily the determinant of a 4×4 matrix is the 4-dimensional volume formed by its 4 column vectors. You can check the formula for the determinant of a 2×2 matrix by computing the area of the parallelogram below using geometry and then comparing that to the value given by the determinant formula.



$$\det\left(\begin{bmatrix} a_1 & a_2 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 4 & 5 \\ 3 & 12 \end{bmatrix}\right) = 48 - 3 \times 5 = 33$$

So now back to our proof, so what happens to the determinant when the column vectors of the matrix $det(A - \lambda I_n) = 0$ are linearly dependent? Remember from the previous notes, this means that either one vector lies in the plane formed by the other two, or all three lie on the same line! In either case, the parallelepiped will be "compressed" into a plane or line, with zero volume. Thus the determinant (volume of the parallelepiped) has to be zero.

Additional Resources For more on determinants, read *Strang* pages 247 - 253 and pages 254 to 257, and try Problem Set 5.1.
In *Schaum's*, read pages 264-265 and pages 265-266, and try Problems 8.1, 8.38, 8.39, 8.2, and 8.40 to 8.43.

9.4 Computing Eigenvalues with Determinants

Now we can go back to finding eigenvalues and eigenvectors. Recall that we are trying to solve the equation

$$(A - \lambda I_n)\vec{x} = \vec{0}$$

Since we are only interested in solutions where $\vec{x} \neq \vec{0}$, we want to find values of λ such that $(A - \lambda I_n)$ has linearly dependent columns. We just learned that a matrix with linearly dependent columns has determinant equal to zero:

$$(A - \lambda I_n)\vec{x} = \vec{0} \implies \det(A - \lambda I_n) = 0 \tag{2}$$

Now, let's use this to find the eigenvalues of a matrix. Consider the example 2×2 matrix

$$A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$$

We now expand the expression above:

$$A - \lambda I_n = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 1-\lambda & 2 \\ 4 & 3-\lambda \end{bmatrix}$$

We now take the determinant of this matrix and set it equal to 0:

$$(1-\lambda)(3-\lambda) - 4 \times 2 = 0$$

Expanding this expression, we get

$$\lambda^2 - 4\lambda - 5 = (\lambda+1)(\lambda-5) = 0$$

We find that there are two solutions to this equation, and therefore two eigenvalues: $\lambda = -1, \lambda = 5$. Each eigenvalue will have its own corresponding eigenvector. To find these, we plug each value of λ into the Equation 2 and solve for $\vec{x}$. Starting with $\lambda = 5$:

$$(A - 5I_2)\vec{x} = \begin{bmatrix} -4 & 2 \\ 4 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \vec{0}$$

We see that both of the rows provide redundant information: $4x_1 - 2x_2 = 0$. The eigenvectors associated with $\lambda = 5$ are all of the form

$$\alpha \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \alpha \in \mathbb{R}$$

Next, we will find the eigenvector associated with $\lambda = -1$. We plug this value of λ into the Equation 2 and solve:

$$(A + I_2)\vec{x} = \begin{bmatrix} 2 & 2 \\ 4 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \vec{0}$$

Both rows provide the information $x_1 = -x_2$. The eigenvectors associated with $\lambda = -1$ are of the form

$$\alpha \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \alpha \in \mathbb{R}$$

In summary, the matrix $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$ has eigenvalues $\lambda = 5, -1$ with associated eigenvectors $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. We can check by multiplying the eignenvectors with A and seeing that they get scaled by their corresponding eigenvalues.

$$\begin{bmatrix}1 & 2\\4 & 3\end{bmatrix}\begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}5\\10\end{bmatrix} = 5\begin{bmatrix}1\\2\end{bmatrix}$$
$$\begin{bmatrix}1 & 2\\4 & 3\end{bmatrix}\begin{bmatrix}1\\-1\end{bmatrix} = \begin{bmatrix}-1\\1\end{bmatrix} = -\begin{bmatrix}1\\-1\end{bmatrix}$$

Additional Resources For more on computing eigenvalues, read *Strang* pages 292 - 296. For additional practice, try Problem Set 6.1.

9.5 Repeated and complex eigenvalues

Every 2×2 matrix has two eigenvalues, but they do not have to be real or unique!

Repeated Eigenvalues:

For a 2×2 matrix, it's possible that the two eigenvalues that you end up with have the same value, leading to a phenomenon called a *Repeated Eigenvalues*. This repeated eigenvalue can one or two associated eigenvectors (unlike a single, unrepeated eigenvalue, which will only have one associated eigenvector). If there are two eigenvectors, they form an *eigenspace*, which is the space of all vectors $\vec{v}$ for which $A\vec{v} = \lambda\vec{v}$.

For example, the following matrix has a repeated eigenvalue of λ.

$$A = \begin{bmatrix}\lambda & 0\\0 & \lambda\end{bmatrix}$$

The *eigenspace* of this matrix is all of $\mathbb{R}^2$ since for any vector $\vec{v} \in \mathbb{R}^2$, $A\vec{v} = \lambda\vec{v}$.

Complex Eigenvalues:

Sometimes when we solve $\det(A - \lambda I_n) = 0$, there will be no real solutions to λ. Consider the transformation that rotates any vector by θ, about the origin:

$$R = \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$$

There are certain specific cases when R has real eigenvalues: When $\theta = 0$, the rotation matrix reduces to the identity matrix, which does not change change vectors, and therefore has eigenvalue $\lambda = 1$. When $\theta = \pi$, the rotation matrix reduces to the negative identity matrix, which changes the sign of vectors and therefore has eigenvalue $\lambda = -1$. In both these cases, all vectors in $\mathbb{R}^2$ are eigenvectors, so $\mathbb{R}^2$ is the eigenspace.

However, when θ does not correspond to 0° or 180° rotation, there are no vectors that are scaled versions of themselves after the transformation. This will results in complex eigenvalues. For example, let's look at

45° rotation:

$$R = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$\det(R - \lambda I) = \frac{1}{2}(1-\lambda)(1-\lambda) + \frac{1}{2}$$

Setting this determinant equal to zero and solving yields the complex eigenvalues, $\lambda = \frac{1}{\sqrt{2}}(1+i)$ and $\lambda = \frac{1}{\sqrt{2}}(1-i)$, which makes sense because there are no physical (real) eigenvectors for this transformation.

9.6 Eigenvectors corresponding to distinct eigenvalues are linearly independent.

Let's explore how the eigenvalues of a 2×2 matrix behave in the most general case. For an arbitrary 2×2 matrix A, we know that its eigenvalues λ will satisfy the equation $\det(A - \lambda I) = 0$. Letting

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

where a, b, c, and d are real scalars, we find that

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

Rearranging, we obtain the quadratic equation

$$\lambda^2 - (a+d)\lambda + (ad - bc) = 0.$$

The polynomial on the left hand side of the above equation is known as the **characteristic polynomial** for the matrix A.

Since the above quadratic equation has real coefficients, we know that there are three possible cases to consider:

1. There are two, real, distinct eigenvalues λ_1 and λ_2 that satisfy the equation.
2. There is a single (repeated) eigenvalue of λ that satisfies the equation.
3. There are two, complex, distinct eigenvalues λ_1 and λ_2 that satisfy the equation.

We won't spend too much time in this course studying the case of complex eigenvalues, so we won't discuss the third case further.

9.6.1 Case 1: Two distinct real eigenvalues

Let's consider the case when we have two distinct, real eigenvalues λ_1 and λ_2 of our matrix A, such that for any $i \in \{1, 2\}$, $\det(A - \lambda_i I) = 0$. Since $\det(A - \lambda_1 I) = 0$, we know that $A - \lambda_1 I$ has a nontrivial nullspace, so there exists a vector $\vec{v}_1 \neq \vec{0}$ such that

$$(A - \lambda I)\vec{v}_1 = \vec{0} \implies A\vec{v}_1 = \lambda_1 \vec{v}_1.$$

As discussed earlier, $\vec{v}_1$ is an eigenvector of A, corresponding to the eigenvalue λ_1. In a similar manner, there will exist an eigenvector $\vec{v}_2$ corresponding to the eigenvalue λ_2, so $A\vec{v}_2 = \lambda_2 \vec{v}_2$.

We will now pose the following question: are $\vec{v}_1$ and $\vec{v}_2$ related in any way? As it turns out, we can show that $\vec{v}_1$ and $\vec{v}_2$ are *linearly independent*, so we cannot write either vector as a linear combination of the other.

Theorem 9.1: Given two eigenvectors $\vec{v}_1$ and $\vec{v}_2$ corresponding to two different eigenvalues λ_1 and λ_2 of a matrix A, it is always the case that $\vec{v}_1$ and $\vec{v}_2$ are linearly independent.

Proof. We will prove this claim by contradiction. Imagine that $\vec{v}_1$ and $\vec{v}_2$ were linearly dependent, so we could express one of them (say $\vec{v}_1$, without loss of generality) as a scalar multiple of the other. Expressed algebraically, we could write $\vec{v}_1 = \alpha \vec{v}_2$ for a suitable scalar α.

Now, since $\vec{v}_1$ and $\vec{v}_2$ are eigenvectors with known eigenvalues, we know by definition that

$$\begin{aligned} A\vec{v}_1 &= \lambda_1 \vec{v}_1 \\ A\vec{v}_2 &= \lambda_2 \vec{v}_2, \end{aligned}$$

where $\lambda_1 \neq \lambda_2$. Substituting in our expression for $\vec{v}_1$ into the first of the above equations, we find that

$$\begin{aligned} & A(\alpha \vec{v}_2) = \lambda_1 (\alpha \vec{v}_2) \\ \implies \quad & \alpha (A\vec{v}_2 - \lambda_1 \vec{v}_2) = \vec{0}. \end{aligned}$$

Since eigenvectors are, by definition, nonzero, we know that $\vec{v}_1 = \alpha \vec{v}_2 \neq 0$, so $\alpha \neq 0$. Thus, we can divide the above equation by α and rearrange to obtain

$$A\vec{v}_2 = \lambda_1 \vec{v}_2.$$

But since $\vec{v}_2$ is an eigenvector with eigenvalue λ_2, we know that

$$A\vec{v}_2 = \lambda_2 \vec{v}_2.$$

Equating the two right hand sides and rearranging, we obtain

$$\lambda_1 \vec{v}_2 = \lambda_2 \vec{v}_2 \implies (\lambda_1 - \lambda_2)\vec{v}_2 = \vec{0}.$$

Finally, since we said that the eigenvalues were distinct, so $\lambda_1 \neq \lambda_2$, it is clear that $\lambda_1 - \lambda_2 \neq 0$. Dividing by this nonzero scalar, we find that

$$\vec{v}_2 = \vec{0}.$$

But $\vec{v}_2$, being an eigenvector, must be nonzero, so this is a contradiction! Our initial hypothesis, that $\vec{v}_1$ and $\vec{v}_2$ were linearly dependent, must therefore be wrong, so the two eigenvectors are in fact linearly independent! This completes the proof. $\square$

9.6.2 Linear independence of eigenvectors with distinct eigenvalues

More generally, it turns out that all eigenvectors with different eigenvalues are linearly independent! More formally, we claim:

Theorem 9.2: Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ be eigenvectors of an $n \times n$ matrix with distinct eigenvalues. It is the case that all the $\vec{v}_i$ are linearly independent from one another. The proof of this theorem is out of scope, but is presented below anyway just for reference for those who are interested.

Proof. This proof will be done by induction over m. To establish some notation, let our $n \times n$ matrix be A, and let our distinct eigenvalues be λ_i, so $A\vec{v}_i = \lambda_i \vec{v}_i$ for all integer $1 \leq i \leq m$.

Now, assume for the sake of induction that the first $k < m$ of our eigenvectors (i.e. $\vec{v}_1, \ldots, \vec{v}_k$) are linearly independent. We will aim to show that the first $k+1$ eigenvectors are also linearly independent. To do so, we will do a proof by contradiction. Imagine, for the sake of contradiction, that the first $k+1$ eigenvectors are in fact linearly dependent, so there exist coefficients $\alpha_1, \ldots, \alpha_{k+1}$ (not all of which are zero) such that

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \ldots + \alpha_k \vec{v}_k + \alpha_{k+1} \vec{v}_{k+1} = \vec{0}.$$

If $\alpha_{k+1} = 0$, then the first k eigenvectors would be linearly dependent, which we know from our inductive hypothesis is not the case. Thus, $\alpha_{k+1} \neq 0$, so we can rearrange the above equation to obtain

$$\vec{v}_{k+1} = -\frac{\alpha_1}{\alpha_{k+1}} \vec{v}_1 - -\frac{\alpha_2}{\alpha_{k+1}} \vec{v}_2 - \ldots - \frac{\alpha_k}{\alpha_{k+1}} \vec{v}_k.$$

For convenience, let $\beta_i = -\alpha_i / \alpha_{k+1}$, so

$$\vec{v}_{k+1} = \beta_1 \vec{v}_1 + \beta_2 \vec{v}_2 + \ldots + \beta_k \vec{v}_k.$$

Notice that, since at least one of the α_i is nonzero, at least one of the β_i is also nonzero.

Now, since $\vec{v}_{k+1}$ is an eigenvector, we have that

$$\begin{aligned} A\vec{v}_{k+1} &= \lambda_{k+1} \vec{v}_{k+1} \\ \Longrightarrow \quad A(\beta_1 \vec{v}_1 + \beta_2 \vec{v}_2 + \ldots + \beta_k \vec{v}_k) &= \lambda_{k+1} (\beta_1 \vec{v}_1 + \beta_2 \vec{v}_2 + \ldots + \beta_k \vec{v}_k) \\ \Longrightarrow \quad \beta_1 (A\vec{v}_1) + \beta_2 (A\vec{v}_2) + \ldots + \beta_k (A\vec{v}_k) &= \beta_1 \lambda_{k+1} \vec{v}_1 + \beta_2 \lambda_{k+1} \vec{v}_2 + \ldots + \beta_k \lambda_{k+1} \vec{v}_k. \end{aligned}$$

But since $\vec{v}_1$ through $\vec{v}_k$ are also eigenvectors where $A\vec{v}_i = \lambda_i \vec{v}_i$, we can substitute to obtain

$$\beta_1 \lambda_1 \vec{v}_1 + \beta_2 \lambda_2 \vec{v}_2 + \ldots + \beta_k \lambda_k \vec{v}_k = \beta_1 \lambda_{k+1} \vec{v}_1 + \beta_2 \lambda_{k+1} \vec{v}_2 + \ldots + \beta_k \lambda_{k+1} \vec{v}_k.$$

Rearranging to pull all the terms on one side, we find that

$$\beta_1 (\lambda_1 - \lambda_{k+1}) \vec{v}_1 + \beta_2 (\lambda_2 - \lambda_{k+1}) \vec{v}_2 + \ldots + \beta_k (\lambda_k - \lambda_{k+1}) \vec{v}_k = \vec{0}.$$

For convenience, let

$$\gamma_i = \beta_i (\lambda_i - \lambda_{k+1}),$$

so

$$\gamma_1 \vec{v}_1 + \gamma_2 \vec{v}_2 + \ldots + \gamma_k \vec{v}_k = \vec{0}.$$

Since all the eigenvalues λ_i are distinct, all the $\lambda_i - \lambda_{k+1} \neq 0$ terms in the above equation are nonzero. Since we showed earlier that at least one of the β_i is nonzero, it is clear that at least one of the γ_i is also nonzero. Thus, we have produced a nontrivial linear combination of the first k eigenvectors that equals zero, so the first k eigenvectors are linearly dependent. But this contradicts our inductive hypothesis, so our initial assumption (that the first $k+1$ eigenvectors were linearly dependent) must be false.

Thus, we have shown that, if at least the first $k < m$ eigenvectors are linearly independent, that the first $k+1$ eigenvectors will also be linearly independent. This will be our induction step.

Clearly, the set containing just the first eigenvector will be linearly independent, since all eigenvectors (including the first) are nonzero. Thus, by induction using the above induction step, we can show that the first m eigenvectors will all be linearly independent, which completes our proof! $\square$

9.6.3 Case 2: A repeated real eigenvalue

In the first case, when we have a single eigenvalue, we know that the corresponding eigenspace may be either one- or two-dimensional. For instance, for

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

we can see that the single eigenvalue is $\lambda = 1$, and the eigenspace is the two dimensional space $\mathbb{R}^2 = \text{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$.

However, for matrices such as

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

we see that the characteristic polynomial

$$\det(A - \lambda I) = \det\left(\begin{bmatrix} -\lambda & 1 \\ 0 & -\lambda \end{bmatrix}\right) = \lambda^2$$

yields the single, repeated eigenvalue $\lambda = 0$. However, computing the associated eigenspace $\text{Null}(A - 0I) = \text{Null}(A)$ yields the one-dimensional eigenspace

$$\text{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}$$

even though the eigenvalue is repeated. In such a case, when the dimension of an eigenspace is less than the number of times the eigenvalue is repeated (called the *multiplicity* of the eigenvalue), we say that the matrix is *defective*. The concept of defective matrices is out of scope for the course, but will be discussed more in EECS 16B. That being said, it is often useful to be aware of their existence, since they can often serve as easy counter-examples if you're not sure if a theorem is true.

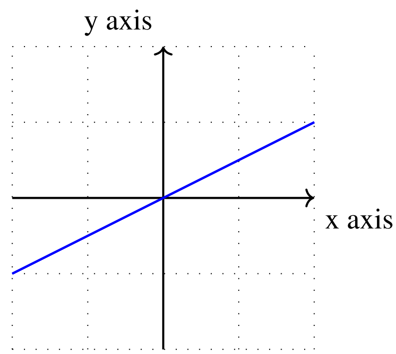
9.7 Examples

In Section 9.4, we saw an example of how to compute the eigenvalues and eigenvectors of a matrix. Here, we'll look at another example where we do the opposite – finding a matrix given it's eigenvalues and eigenvectors. Finally, we'll go back to our page rank example and rank the popularity of webpages by finding the appropriate eigenvector.

Example 9.1 (Finding a matrix from its eigenvalues and eigenvectors): Now we turn to the case where we try to derive the matrix from the provided vectors. This is essentially reverse engineering the previous problem and reworking the derivation of the eigenvalues.

We would like to solve the problem: what is the matrix A that reflects any $2D$ vectors over the vector $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$?

Below is a graphical representation of that vector:



Remember that an eigenvector of a matrix is any vector that results in a scaled version of itself after being transformed by the matrix. Can you think of any eigenvectors of this reflection transformation? First we observe that anything lying on the line of reflection will be unchanged, in other words, scaled by one. Therefore, we can write the eigenvector, eigenvalue pair:

$$\lambda_1 = 1$$

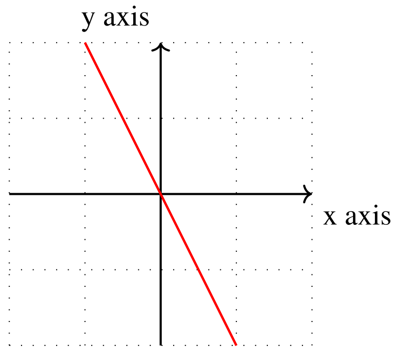
$$\vec{v}_1 = \alpha \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

In addition, any vector lying perpendicular to the line of reflection will be transformed to point in the *opposite* direction, in other words, scaled by -1. This results in another eigenvector, eigenvalue pair:

$$\lambda_2 = -1$$

$$\vec{v}_2 = \alpha \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

Now we can use our definition, $A\vec{v} = \lambda\vec{v}$, to set up a system of equations so we can solve for matrix A:



$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

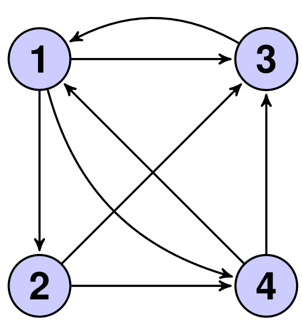
We rearrange these linear equations as follows. (*To see these are equivalent, try writing out all four equations in terms of* $a_1 \ldots a_4$.)

$$\begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 1 \\ -2 \end{bmatrix}$$

Finally, we solve the system of equations with Gaussian Elimination to get our transformation matrix, A.

$$A = \begin{bmatrix} .6 & .8 \\ .8 & -.6 \end{bmatrix}$$

Example 9.2 (Page Rank): Let's go back to our page rank example from the beginning of this note. Recall that we have four webpages, with links between them represented by the diagram below:



We can write the tranformation matrix P that tells us how the number of people on each webpage changes over time:

$$\vec{x}(t+1) \;=\; P\vec{x}(t) \;=\; \begin{bmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{bmatrix} \vec{x}(t)$$

where $\vec{x}(t)$ is a vector describing the fraction of people on each page at time t.

We are interested in the *steady state* vector, $\vec{x^*}$ such that

$$\vec{x^*} = P\vec{x^*}.$$

Looking back at our definition of eigenvalues and eigenvectors, we can see that $\vec{x^*}$ is the eigenvector associated with eigenvalue $\lambda = 1$. In this specific case of finding a steady state vector, we already know the eigenvalue, so proceed to finding the associated eigenvector. (More generally, we would use the determinant to find the eigenvalues first).

From Equation 1, we can write that our steady state vector satisfies

$$(P - \lambda I)\vec{x^*} = \vec{0}$$

where $\lambda = 1$.

$$\left(\begin{bmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \right) \vec{x^*} = \vec{0}$$

$$\begin{bmatrix} -1 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & -1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & -1 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & -1 \end{bmatrix} \vec{x^*} = \vec{0}$$

We solve for $\vec{x^*}$ by finding the null space of the above matrix. After performing Gaussian elimination on the matrix above, we get the following row equivalent system of equations:

$$\begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -\frac{2}{3} \\ 0 & 0 & 1 & -\frac{3}{2} \\ 0 & 0 & 0 & 0 \end{bmatrix} \vec{x^*} = \vec{0}$$

Remember that each row represents an equation:

$$x_1 - 2x_4 = 0$$
$$x_2 - \frac{2}{3}x_4 = 0$$
$$x_3 - \frac{3}{2}x_4 = 0$$

We can represent this set of solutions as a vector where x_4 is a free variable, $x_4 \in \mathbb{R}$.

$$\vec{x^*} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2x_4 \\ \frac{2}{3}x_4 \\ \frac{3}{2}x_4 \\ x_4 \end{bmatrix}$$

Since we want, the steady state vector to represent the fraction of people on each page, we can choose x_4 so that the total number of people sums to 1.

$$\vec{x^*} = \frac{1}{31} \begin{bmatrix} 12 \\ 4 \\ 9 \\ 6 \end{bmatrix}$$

From our steady state solution, we conclude that webpage 1 is most popular, then webpages 3, 4, and 2.

9.8 Steady States

We know that a steady state $\vec{x}^*$ of a transformation matrix A is defined to be such that

$$A\vec{x}^* = \vec{x}^*.$$

In other words, it is an element of the eigenspace of A corresponding to the eigenvalue $\lambda = 1$. The above equation tells us that if we start at a steady state, then we will remain unaffected by the transformation matrix over time.

However, what happens if we *don't* start at a steady state? Can we say anything about what happens to our state as we keep applying the transformation matrix?

9.8.1 Starting at eigenvectors

We will first look at a simpler problem - what happens if our initial state is an eigenvector $\vec{v}$ of A, but with a real eigenvalue λ that may not equal 1? In other words, $A\vec{v} = \lambda\vec{v}$, but it is possible that $\lambda\vec{v} \neq \vec{v}$. Let's try to see what happens over time, as we keep applying A to our initial state $\vec{x}[0] = \vec{v}$. For the first few timesteps, we see that

$$\begin{aligned} & \vec{x}[0] = \vec{v} \\ \implies \quad & \vec{x}[1] = A\vec{x}[0] = A\vec{v} = \lambda\vec{v} \\ \implies \quad & \vec{x}[2] = A\vec{x}[1] = A(\lambda\vec{v}) = \lambda^2\vec{v} \\ \implies \quad & \vec{x}[3] = A\vec{x}[2] = A(\lambda^2\vec{v}) = \lambda^3\vec{v} \\ & \vdots \end{aligned}$$

We can see, therefore, that after t timesteps, our state will become

$$\vec{x}[t] = \lambda^t \vec{v}.$$

But how does this expression behave as time continues, and $n \to \infty$? We can see that there are a number of cases to consider:

- If $\lambda > 1$, then it can be seen that $\lambda, \lambda^2, \lambda^3, \ldots$ will keep increasing to $+\infty$, so our state $\vec{x}[t]$ will keep growing along the direction of $\vec{v}$ towards infinity.
- If $\lambda = 1$, we're just at the steady state, so $\vec{x}[t] = \vec{v}$ as $t \to \infty$.
- If $0 < \lambda < 1$, then $\lambda, \lambda^2, \lambda^3, \ldots$ will decrease towards zero from above, so our state $\vec{x}[t]$ will keep scaling down and approach $\vec{0}$ in the limit.
- If $\lambda = 0$, then $\vec{x}[1] = 0\vec{v} = \vec{0}$, so our state will immediately drop to $\vec{0}$ from the first timestep onwards.
- If $\lambda < 0$, then behavior similar to the above will occur based on the value of $|\lambda|$, except that the sign of the state will keep flipping at each step. For instance, if $\lambda < 0$ but $|\lambda| > 1$ (in other words, $\lambda < -1$), then the magnitude of the state will tend towards infinity but will keep flipping its sign each time step.

9.8.2 General initial states

Still, we're not done! All we've considered so far are steady states of the form $\vec{x}[0] = \vec{v}$ where $\vec{v}$ is an eigenvector of the transformation matrix A. What if we have an arbitrary initial state, which in general may not be an eigenvector of A?

To deal with this issue, we will make our lives a little easier by considering only matrices whose eigenvectors form a basis (i.e.span all of the state space) — in other words, if A is $n \times n$, then *any* initial state $\vec{x}[0]$ can be written as a linear combination

$$\vec{x}[0] = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \ldots + \alpha_n \vec{v}_n,$$

where $\vec{v}_1$ through $\vec{v}_n$ are linearly independent eigenvectors with eigenvalues λ_1 through λ_n, and the α_i are real scalar coefficients. The general case, when this property is not guaranteed to hold, will be discussed in EECS 16B.

Now, we have previously seen that

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

We will substitute our linear combination for $\vec{x}[0]$ in terms of the eigenvectors into the above equation, to obtain

$$\begin{aligned}
\vec{x}[t] &= A^t(\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \ldots + \alpha_n \vec{v}_n) \\
&= \alpha_1(A^t \vec{v}_1) + \alpha_2(A^t \vec{v}_2) + \ldots + \alpha_n(A^t \vec{v}_n) \\
&= \alpha_1(\lambda_1^t \vec{v}_1) + \alpha_2(\lambda_2^t \vec{v}_2) + \ldots + \alpha_n(\lambda_n^t \vec{v}_n).
\end{aligned}$$

Thus, to determine the behavior of an (almost) arbitrary linear system with an arbitrary initial state, all we have to do is write the initial state as a linear combination of the eigenvectors, see how each term of the linear combination behaves as $t \to \infty$, and plug back into the linear combination to determine how the initial state evolves as $t \to \infty$.

We have expressed $\vec{x}[t]$ as a linear combination of the $\lambda_i^t \vec{v}_i$, which each represent what happens to each eigenvector after repeated applications of A. And, from the previous section, we know how this behavior depends on the value of λ_i!

9.8.3 Computing linear combinations

In the above description we wrote out $\vec{x}[0]$ as a linear combination of the eigenvectors. In general, how do we write an arbitrary initial state as a linear combination of n linearly independent eigenvectors $\{\vec{v}_i\}$? We'd like to solve for the coefficients $\{\alpha_i\}$ in the equation

$$\vec{x}[0] = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \ldots + \alpha_n \vec{v}_n.$$

Rewriting it as a matrix multiplication using the linear combination interpretation and stacking the unknown α_i into a vector, we obtain the linear system

$$\begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \vec{x}[0],$$

which we can solve using Gaussian elimination. Alternatively, since the matrix on the left is $n \times n$ with n linearly independent columns, we can pre-multiply by its inverse to express our solution vector as

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}^{-1} \vec{x}[0].$$

We will revisit this in EECS16B when we talk about the diagonalization.

9.8.4 Criteria for convergence

Now, we have all the tools needed to determine what happens to an initial state after repeated application of a transformation matrix.

Consider an arbitrary initial state $\vec{x}[0]$. We know that

$$\vec{x}[t] = \alpha_1(\lambda_1^t \vec{v}_1) + \alpha_2(\lambda_2^t \vec{v}_2) + \ldots + \alpha_n(\lambda_n^t \vec{v}_n).$$

Essentially, we want to know if $\vec{x}[t]$ will ultimately converge to some fixed value, *no matter* what the $\{\alpha_i\}$ are.

Observe that if a single λ_i is such that $|\lambda_i| > 1$, that $\lambda_i^t \vec{v}_i \to \infty$ in the limit. Thus, our state $\vec{x}[t]$ could potentially go to infinity in the limit. Furthermore, observe that if a single $\lambda_i = -1$, that $\lambda_i^t \vec{v}_i$ will oscillate back and forth in the limit, neither decreasing towards zero or growing to infinity. Thus, our state $\vec{x}[t]$ could potentially oscillate indefinitely, and so fail to converge to a fixed value.

However, if all the λ_i are such that $-1 < \lambda_i \leq 1$, then each term in the linear combination will either go to zero (if the corresponding $\lambda_i \neq 1$) or stay the same (if $\lambda_i = 1$), meaning that the overall state $\vec{x}[t]$ will always converge to a fixed value, no matter what the initial coefficients are.

Consequently, we have come up with a condition that is sufficient to ensure that any initial state of a transformation matrix will reach a steady state in the limit. But to compute what exactly this steady state is, or to determine if a steady state is reached for a *particular* initial state even if it will not always be reached in the general case, we need to go back to our technique of writing our initial state as a linear combination of eigenvectors and computing the limit of each term in the linear combination.

9.9 Practice Problems

These practice problems are also available in an interactive form on the course website.

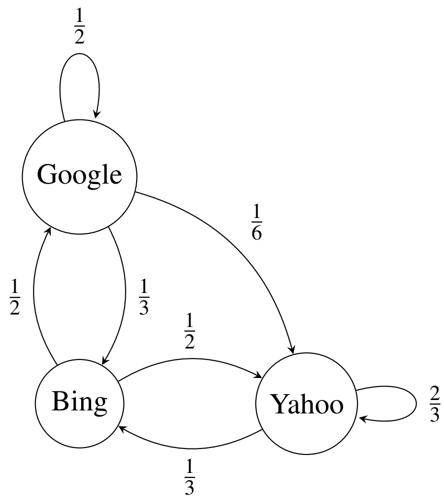
1. What are the eigenvalues of the matrix $\begin{bmatrix} 3 & 2 \\ 4 & 1 \end{bmatrix}$?
2. What are the eigenvectors of the matrix $\begin{bmatrix} 3 & 2 \\ 4 & 1 \end{bmatrix}$?
3. True or False: If an $n \times n$ matrix $\mathbf{A}$ is not invertible, then it has an eigenvalue $\lambda = 0$.
4. True or False: If an invertible matrix $\mathbf{A}$ has an eigenvalue λ, then $\mathbf{A}^{-1}$ has the eigenvalue $\frac{1}{\lambda}$.
   (a) Always true.
   (b) False.
   (c) True only if $\lambda \neq 0$.
5. Two students find two different eigenvectors associated with the same eigenvalue of a matrix, $\begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -\frac{1}{2} \\ 2 \end{bmatrix}$. Student One insists that the other must have made an error because every eigenvalue is associated with a unique eigenvector. Student Two insists that they are both valid answers. Who is right, Student One or Student Two?
6. True or False: If for some matrix $\mathbf{A}$, $\mathbf{A}\vec{x}_1 = \lambda_1\vec{x}_1$ and $\mathbf{A}\vec{x}_1 = \lambda_2\vec{x}_1$, then $\lambda_1 = \lambda_2$.
7. True or False: A matrix with only real entries can have complex eigenvalues.
8. True or False: If $\mathbf{A}^T$ has an eigenvalue λ, then $\mathbf{A}$ also has the eigenvalue λ.
9. True or False: A diagonal $n \times n$ matrix has n distinct eigenvalues.
10. For a square non-invertible $n \times n$ matrix $\mathbf{A}$, what is the maximum and minimum number of distinct eigenvalues it can have?
    (a) Max: 1, Min: 0
    (b) Max: n, Min: 1

(c) Max: n, Min: $n-1$

(d) Max: $n-1$, Min: 0

(e) Max: $n-1$, Min: 1

11. Find the steady state, if it exists, of the following system assuming that the state vector is $\vec{x}[t] = \begin{bmatrix} x_{\text{Google}}[t] \\ x_{\text{Yahoo}}[t] \\ x_{\text{Bing}}[t] \end{bmatrix}$ and that we start off with $\vec{x}[0] = \begin{bmatrix} 100 \\ 100 \\ 100 \end{bmatrix}$.



12. What is the matrix $\mathbf{A}$ that reflects any vector in $\mathbb{R}^2$ about the line through the origin in the direction of $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$?