

Key Concepts

By the end of this note, you should be able to do the following:

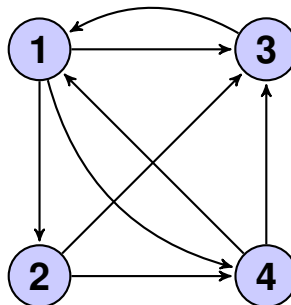
1. Understand conceptually what Eigenvalues and Eigenvectors are, especially in the context of Google's Page Rank algorithm.
2. Understand how to find the Eigenvalues and Eigenvectors of a given matrix.
3. Use the Eigenvalues and Eigenvectors of an unknown matrix to derive its components.

Introduction: Page Rank

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, not everyone visits outgoing links at once, browsing sessions vary in length, and people often seek out entirely new pages over clicking outgoing links from their current page. 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:



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 a handy matrix notation. Define $\vec{x}(k) = [x_1(k) \ x_2(k) \ \dots \ x_n(k)]^T$ 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

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

We usually think of $\vec{x}(k)$ as fractions of viewers, so we initialize $\vec{x}(0)$ as $[\frac{1}{4} \ \frac{1}{4} \ \frac{1}{4} \ \frac{1}{4}]^T$. 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.

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

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

Eigenvectors and Eigenvalues

In our Page Rank example, \vec{x}^* is an *eigenvector* of P . An $n \times 1$ vector \vec{x} is called an eigenvector of an $n \times n$ square matrix A if

$$A\vec{x} = \lambda\vec{x}, \vec{x} \neq \vec{0}, \lambda \in \mathbb{R}$$

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:

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

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 rewrite the product in terms of the columns of $A - \lambda I_n$:

$$\begin{bmatrix} | & | & \dots & | \\ \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} | \\ \vec{a}_1 \\ | \end{bmatrix} v_1 + \begin{bmatrix} | \\ \vec{a}_2 \\ | \end{bmatrix} v_2 + \dots + \begin{bmatrix} | \\ \vec{a}_n \\ | \end{bmatrix} v_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

(Check to see that this column formulation is a valid way to represent matrix multiplication). From this formulation, we can see that some nonzero linear combination of the columns of $A - \lambda I_n$ results in 0. This means that the columns of this matrix must be linearly dependent. Now we need a way to compute the value of λ . To do this, we will introduce the determinant.

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.

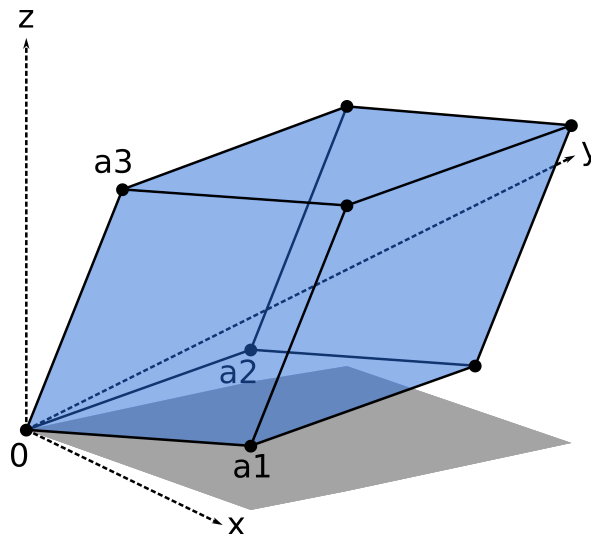
For this class, we only need to know how to compute the determinant of a 2×2 matrix by hand. This is given by

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

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.

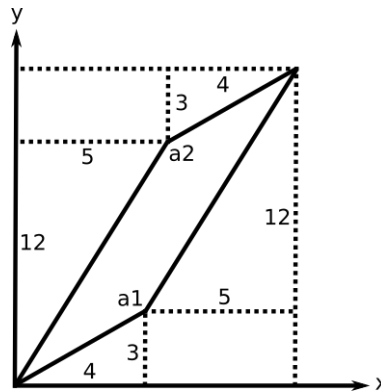
There is a beautiful geometrical connection between the column vectors of a matrix and its determinant. For simplicity, first consider a 3×3 matrix. Let its column vectors be a_1 , a_2 and a_3 . Then consider the parallelepiped formed by these vectors:¹



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.

¹Modified from image by Claudio Rocchini, <https://en.wikipedia.org/wiki/Determinant>

Similarly 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.



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

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.

Computing Eigenvalues with Determinants

With this, we can revisit our previous equation:

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

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 eigenvalues: $\lambda = -1$, $\lambda = 5$. Each eigenvalue will have its own corresponding eigenvector. To find these, we simply plug in the values of λ into the original equation:

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

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

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

Now we plug in $\lambda = -1$:

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

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

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

Eigenvalue/Eigenvector Vocabulary

Just a reminder that the general formula for finding the eigenvector of a $n \times n$ matrix is:

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

where \vec{v} is the $n \times 1$ eigenvector, and λ is the scalar eigenvalue.

From that relationship we are also reminded that:

$$\det(\lambda I - A) = 0$$

where λ is an eigenvalue of A .

Note that every 2×2 linear transformation has two complex eigenvectors, but it can have anywhere between zero to two real eigenvectors, so they are not guaranteed to be real in any case!

Repeated Eigenvalues:

For a 2×2 matrix, it's possible that the two eigenvalues that you end up with are the same one, leading to a phenomenon called a *Repeated Eigenvalues*. When there is only one eigenvalue, the eigenvalue is called a repeated eigenvalue. You can have one or two eigenvectors, and if there is two they form an eigenspace.

For example, the following matrix has a repeated eigenvalue:

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

Rotation Matrices Theoretical and Practical:

Consider the transformation that rotates any vector by θ , about the origin. Then, there are eigenvectors only for particular values of θ : 0 , π , and the corresponding eigenvalues are 1 and -1 respectively. It is clear that, for these transformations, any vector is an eigenvector. Here we can choose any two vectors to refer as eigenvectors and we'll have them pulled from the eigenvalues. The space defined by these two vectors (R-Squared) is called the eigenspace. If θ is any other value, you run the risk of them being complex.

Further Examples

We're going to take two example problems to outline how to approach different situations. The first one is where we talk a matrix and find the eigenvectors and eigenvalues. Let's take the following matrix A :

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

We would like to find the eigenvectors and eigenvalues possible for matrix A . To do this, we start with the initial equation for the eigenvector which is:

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

where \vec{v} is not equal to 0 . Then we apply the determinant process to root out the eigenvalues!

$$(\lambda I - A)\vec{v} = 0$$

$$\det(\lambda I - A) = 0$$

$$\lambda_1 = 5$$

$$\lambda_2 = -1$$

Therefore, we find the eigenvectors to be the following:

$$\vec{v}_1 = \alpha \begin{bmatrix} .5 \\ 1 \end{bmatrix}, \alpha \in \mathbb{R}$$

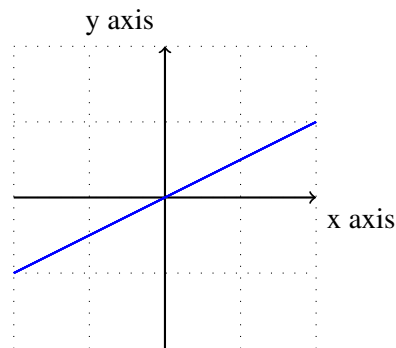
$$\vec{v}_2 = \alpha \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \alpha \in \mathbb{R}$$

Finding Matrix from 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 action matrix A that reflect any $2D$ vectors over the vector $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$?

Below is a graphical representation of that vector:



First, observe that anything lying on that line has the eigenvector and eigenvalue

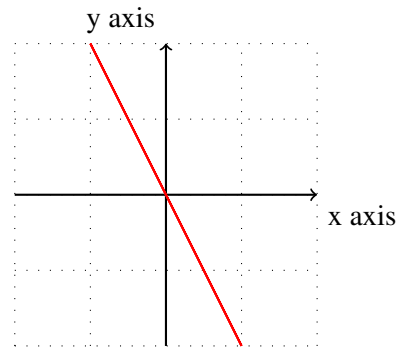
$$\lambda_1 = 1$$

$$v_1 = \alpha \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

In addition, for anything lying on the line orthogonal to the original line, we have

$$\lambda_2 = -1$$

$$v_2 = \alpha \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$



Since we now have the eigenvectors and eigenvalues, using the equation $A\vec{v} = \lambda\vec{v}$, we can solve for matrix [A] by solving the following equations:

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

We find that:

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