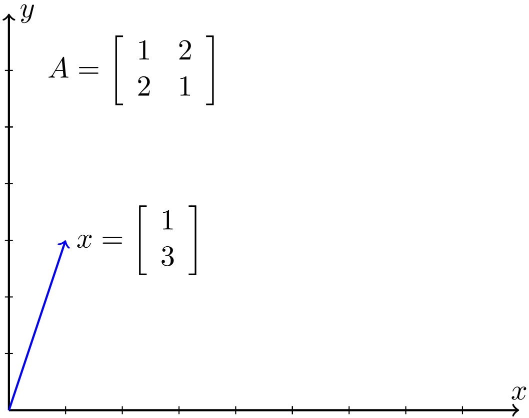
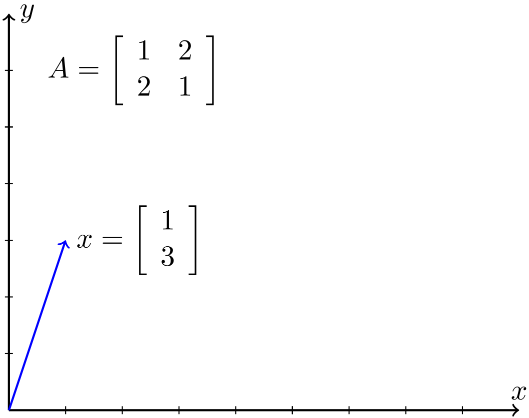
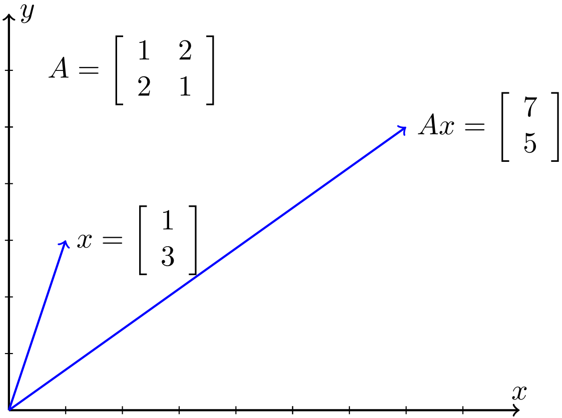
Module 6.1 : Eigenvalues and Eigenvectors



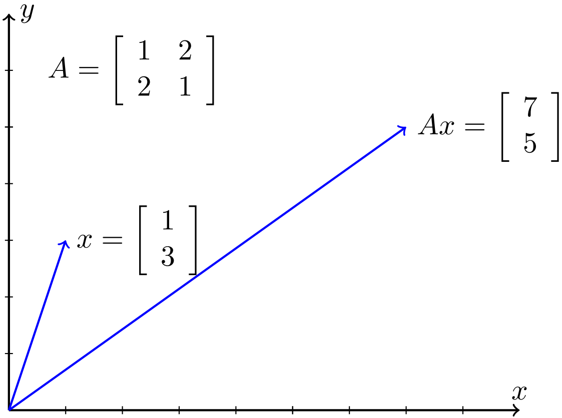
- What happens when a matrix hits a vector?



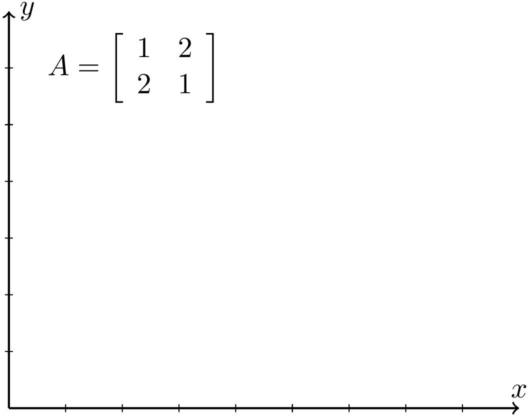
- What happens when a matrix hits a vector?
- The vector gets transformed into a new vector (it strays from its path)



- What happens when a matrix hits a vector?
- The vector gets transformed into a new vector (it strays from its path)

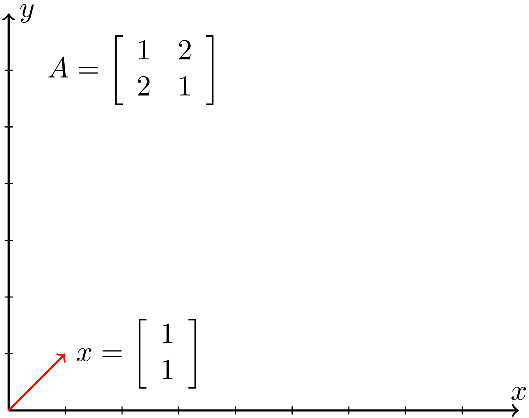


- What happens when a matrix hits a vector?
- The vector gets transformed into a new vector (it strays from its path)
- The vector may also get scaled (elongated or shortened) in the process.

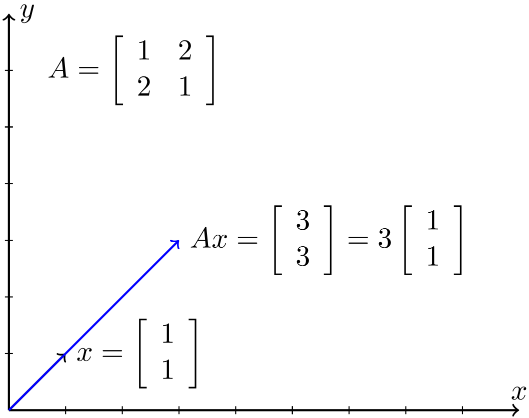


$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

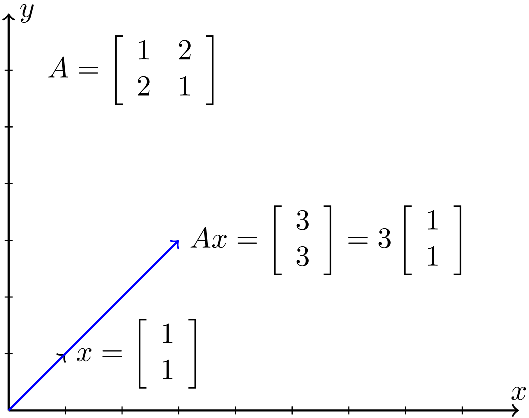
- For a given square matrix A, there exist special vectors which refuse to stray from their path.



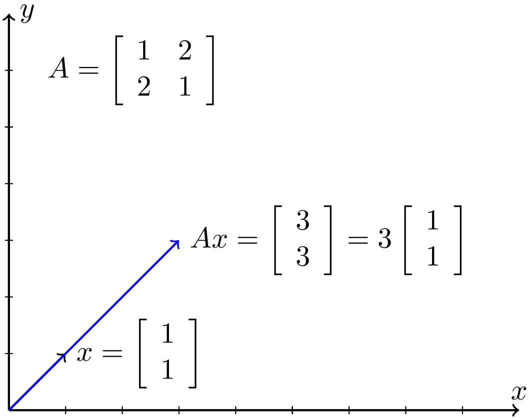
- For a given square matrix A, there exist special vectors which refuse to stray from their path.



- For a given square matrix A, there exist special vectors which refuse to stray from their path.

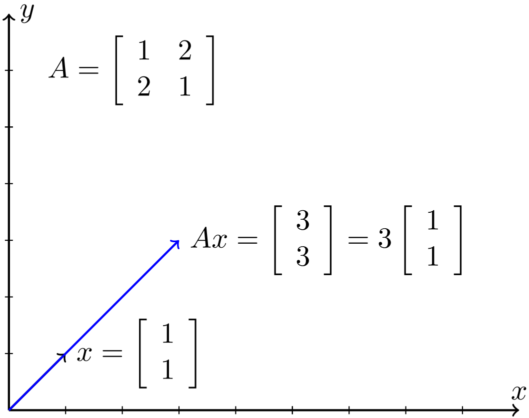


- For a given square matrix A, there exist special vectors which refuse to stray from their path.
- These vectors are called eigenvectors.



- For a given square matrix A, there exist special vectors which refuse to stray from their path.
- These vectors are called eigenvectors.
- More formally,

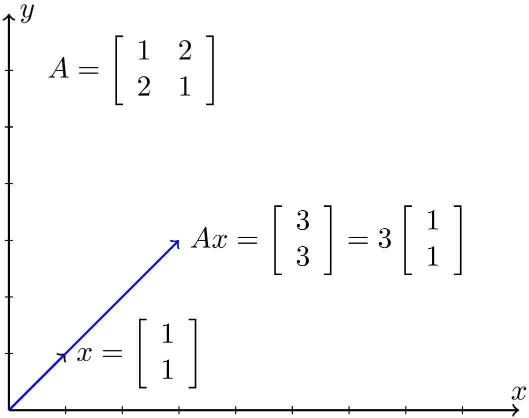
  $Ax = \lambda x$ [direction remains the same]

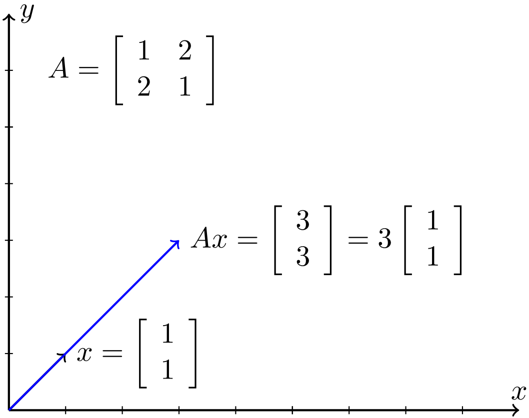


- For a given square matrix A, there exist special vectors which refuse to stray from their path.
- These vectors are called eigenvectors.
- More formally,

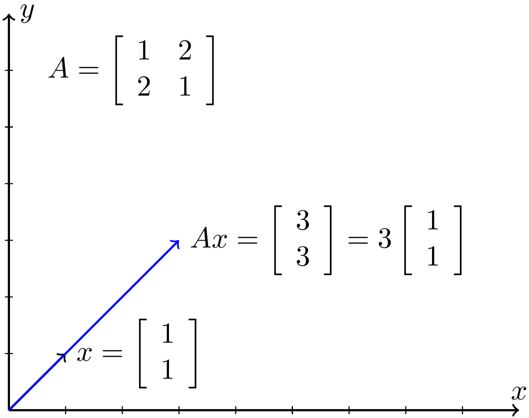
$$Ax = \lambda x \text{ [direction remains the same]}$$

- The vector will only get scaled but will not change its direction.

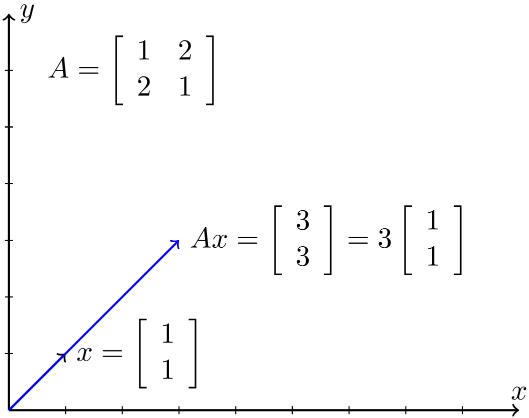

y
$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$
$Ax = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
$x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
x



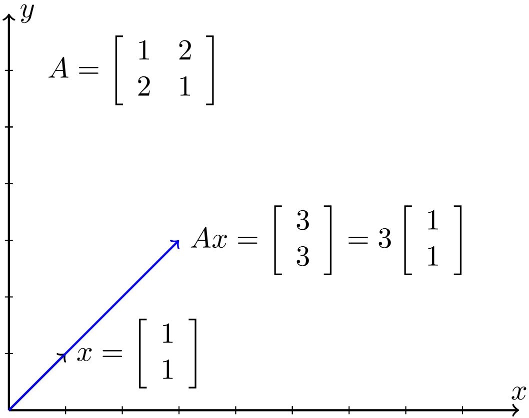
- So what is so special about eigenvectors?



- So what is so special about eigenvectors?
- Why are they always in the limelight?

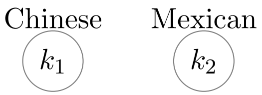


- So what is so special about eigenvectors?
- Why are they always in the limelight?
- It turns out that several properties of matrices can be analyzed based on their eigenvalues (for example, see spectral graph theory)



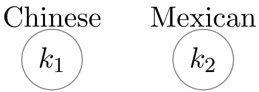
- So what is so special about eigenvectors?
- Why are they always in the limelight?
- It turns out that several properties of matrices can be analyzed based on their eigenvalues (for example, see spectral graph theory)
- We will now see two cases where eigenvalues/vectors will help us in this course

- Let us assume that on day 0, k_1 students eat Chinese food, and k_2 students eat Mexican food. (Of course, no one eats in the mess!)



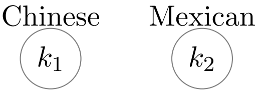
$$v_{(0)} = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$$

- Let us assume that on day 0, k_1 students eat Chinese food, and k_2 students eat Mexican food. (Of course, no one eats in the mess!)



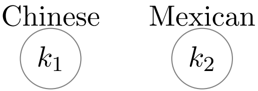
$$v_{(0)} = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$$

- Let us assume that on day 0, k_1 students eat Chinese food, and k_2 students eat Mexican food. (Of course, no one eats in the mess!)
- On each subsequent day i, a fraction p of the students who ate Chinese food on day $(i-1)$, continue to eat Chinese food on day i, and $(1-p)$ shift to Mexican food.



$$v_{(0)} = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$$

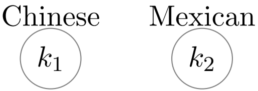
- Let us assume that on day 0, k_1 students eat Chinese food, and k_2 students eat Mexican food. (Of course, no one eats in the mess!)
- On each subsequent day i, a fraction p of the students who ate Chinese food on day $(i-1)$, continue to eat Chinese food on day i, and $(1-p)$ shift to Mexican food.
- Similarly a fraction q of students who ate Mexican food on day $(i-1)$ continue to eat Mexican food on day i, and $(1-q)$ shift to Chinese food.



$$v_{(0)} = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$$

$$v_{(1)} = \begin{bmatrix} pk_1 + (1-q)k_2 \\ (1-p)k_1 + qk_2 \end{bmatrix}$$

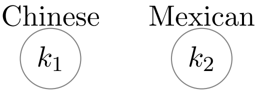
- Let us assume that on day 0, k_1 students eat Chinese food, and k_2 students eat Mexican food. (Of course, no one eats in the mess!)
- On each subsequent day i, a fraction p of the students who ate Chinese food on day $(i-1)$, continue to eat Chinese food on day i, and $(1-p)$ shift to Mexican food.
- Similarly a fraction q of students who ate Mexican food on day $(i-1)$ continue to eat Mexican food on day i, and $(1-q)$ shift to Chinese food.



$$v_{(0)} = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$$

$$v_{(1)} = \begin{bmatrix} pk_1 + (1-q)k_2 \\ (1-p)k_1 + qk_2 \end{bmatrix} = \begin{bmatrix} p & 1-q \\ 1-p & q \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$$

- Let us assume that on day 0, k_1 students eat Chinese food, and k_2 students eat Mexican food. (Of course, no one eats in the mess!)
- On each subsequent day i, a fraction p of the students who ate Chinese food on day $(i-1)$, continue to eat Chinese food on day i, and $(1-p)$ shift to Mexican food.
- Similarly a fraction q of students who ate Mexican food on day $(i-1)$ continue to eat Mexican food on day i, and $(1-q)$ shift to Chinese food.

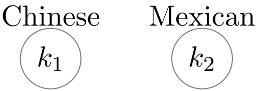


$$v_{(0)} = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$$

$$v_{(1)} = \begin{bmatrix} pk_1 + (1-q)k_2 \\ (1-p)k_1 + qk_2 \end{bmatrix}$$
$$= \begin{bmatrix} p & 1-q \\ 1-p & q \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$$

$$\begin{aligned} v_{(1)} &= Mv_{(0)} \\ v_{(2)} &= Mv_{(1)} \\ &= M^2 v_{(0)} \end{aligned}$$

- Let us assume that on day 0, k_1 students eat Chinese food, and k_2 students eat Mexican food. (Of course, no one eats in the mess!)
- On each subsequent day i, a fraction p of the students who ate Chinese food on day $(i-1)$, continue to eat Chinese food on day i, and $(1-p)$ shift to Mexican food.
- Similarly a fraction q of students who ate Mexican food on day $(i-1)$ continue to eat Mexican food on day i, and $(1-q)$ shift to Chinese food.



$$v_{(0)} = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$$

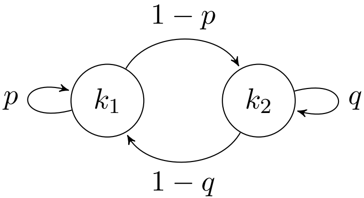
$$\begin{aligned} v_{(1)} &= \begin{bmatrix} pk_1 + (1-q)k_2 \\ (1-p)k_1 + qk_2 \end{bmatrix} \\ &= \begin{bmatrix} p & 1-q \\ 1-p & q \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} v_{(1)} &= Mv_{(0)} \\ v_{(2)} &= Mv_{(1)} \\ &= M^2 v_{(0)} \end{aligned}$$

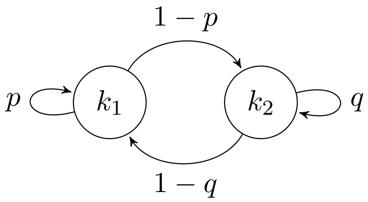
In general, $v_{(n)} = M^n v_{(0)}$

- Let us assume that on day 0, k_1 students eat Chinese food, and k_2 students eat Mexican food. (Of course, no one eats in the mess!)
- On each subsequent day i, a fraction p of the students who ate Chinese food on day $(i-1)$, continue to eat Chinese food on day i, and $(1-p)$ shift to Mexican food.
- Similarly a fraction q of students who ate Mexican food on day $(i-1)$ continue to eat Mexican food on day i, and $(1-q)$ shift to Chinese food.
- The number of customers in the two restaurants is thus given by the following series:

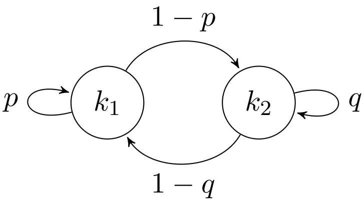
$$v_{(0)}, Mv_{(0)}, M^2 v_{(0)}, M^3 v_{(0)}, \ldots$$

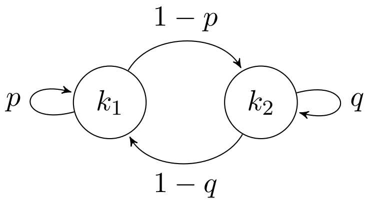

$1 - p$
p
k_1
k_2
q
$1 - q$

- This is a problem for the two restaurant owners.

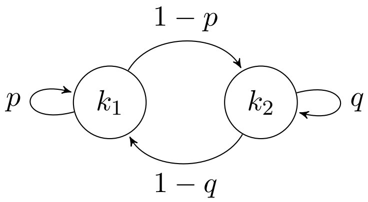


- This is a problem for the two restaurant owners.
- The number of patrons is changing constantly.

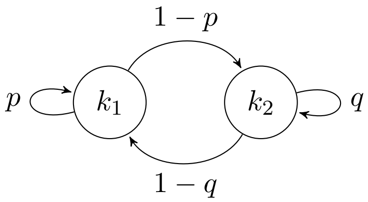




- This is a problem for the two restaurant owners.
- The number of patrons is changing constantly.
- Or is it? Will the system eventually reach a steady state? (i.e. will the number of customers in the two restaurants become constant over time?)



- This is a problem for the two restaurant owners.
- The number of patrons is changing constantly.
- Or is it? Will the system eventually reach a steady state? (i.e. will the number of customers in the two restaurants become constant over time?)
- Turns out they will!



- This is a problem for the two restaurant owners.
- The number of patrons is changing constantly.
- Or is it? Will the system eventually reach a steady state? (i.e. will the number of customers in the two restaurants become constant over time?)
- Turns out they will!
- Let's see how?

Definition

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvectors of an $n \times n$ matrix A. λ_1 is called the dominant eigen value of A if

$$|\lambda_1| \geq |\lambda_i| \; i = 2, \ldots, n$$

Definition

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvectors of an $n \times n$ matrix A. λ_1 is called the dominant eigen value of A if

$$|\lambda_1| \geq |\lambda_i| \; i = 2, \ldots, n$$

Definition

A matrix M is called a stochastic matrix if all the entries are positive and the sum of the elements in each column is equal to 1.
(Note that the matrix in our example is a stochastic matrix)

Definition

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvectors of an $n \times n$ matrix A. λ_1 is called the dominant eigen value of A if

$$|\lambda_1| \geq |\lambda_i| \; i = 2, \ldots, n$$

Definition

A matrix M is called a stochastic matrix if all the entries are positive and the sum of the elements in each column is equal to 1.
(Note that the matrix in our example is a stochastic matrix)

Theorem

The largest (dominant) eigenvalue of a stochastic matrix is 1.
See proof here

Definition

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvectors of an $n \times n$ matrix A. λ_1 is called the dominant eigen value of A if

$$|\lambda_1| \geq |\lambda_i| \; i = 2, \ldots, n$$

Theorem

The largest (dominant) eigenvalue of a stochastic matrix is 1.
See proof here

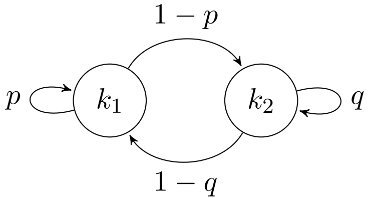
Definition

A matrix M is called a stochastic matrix if all the entries are positive and the sum of the elements in each column is equal to 1.
(Note that the matrix in our example is a stochastic matrix)

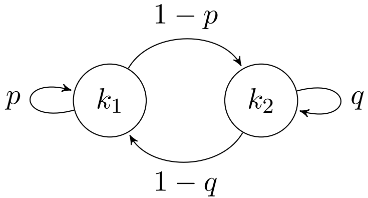
Theorem

If A is a $n \times n$ square matrix with a dominant eigenvalue, then the sequence of vectors given by Av_0, A^2v_0, $\ldots$, $A^n v_0, \ldots$ approaches a multiple of the dominant eigenvector of A.
(the theorem is slightly misstated here for ease of explanation)

- Let e_d be the dominant eigenvector of M and $\lambda_d = 1$ the corresponding dominant eigenvalue

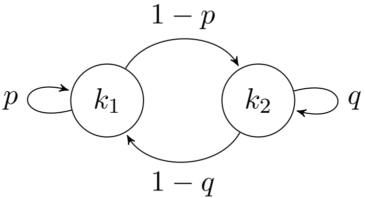


- Let e_d be the dominant eigenvector of M and $\lambda_d = 1$ the corresponding dominant eigenvalue
- Given the previous definitions and theorems, what can you say about the sequence $Mv_{(0)}, M^2 v_{(0)}, M^3 v_{(0)}, \ldots$?



- Let e_d be the dominant eigenvector of M and $\lambda_d = 1$ the corresponding dominant eigenvalue
- Given the previous definitions and theorems, what can you say about the sequence $Mv_{(0)}, M^2v_{(0)}, M^3v_{(0)}, \ldots$?
- There exists an n such that

$$v_{(n)} = M^n v_{(0)} = ke_d \text{ (some multiple of } e_d)$$

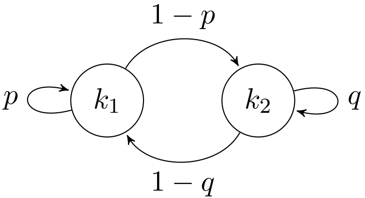


- Let e_d be the dominant eigenvector of M and $\lambda_d = 1$ the corresponding dominant eigenvalue
- Given the previous definitions and theorems, what can you say about the sequence $Mv_{(0)}, M^2v_{(0)}, M^3v_{(0)}, \ldots$?
- There exists an n such that

$$v_{(n)} = M^n v_{(0)} = ke_d \text{ (some multiple of } e_d\text{)}$$

- Now what happens at time step $(n+1)$?

$$v_{(n+1)} = Mv_{(n)} = M(ke_d) = k(Me_d) = k(\lambda_d e_d) = ke_d$$



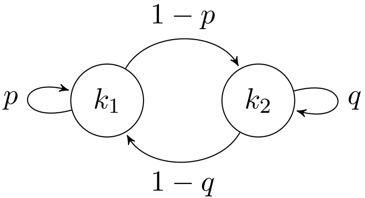
- Let e_d be the dominant eigenvector of M and $\lambda_d = 1$ the corresponding dominant eigenvalue
- Given the previous definitions and theorems, what can you say about the sequence $Mv_{(0)}, M^2v_{(0)}, M^3v_{(0)}, \ldots$?
- There exists an n such that

$$v_{(n)} = M^n v_{(0)} = ke_d \text{ (some multiple of } e_d)$$

- Now what happens at time step $(n+1)$?

$$v_{(n+1)} = Mv_{(n)} = M(ke_d) = k(Me_d) = k(\lambda_d e_d) = ke_d$$

- The population in the two restaurants becomes constant after time step n.
  See Proof Here



- Now instead of a stochastic matrix let us consider any square matrix A

- Now instead of a stochastic matrix let us consider any square matrix A
- Let p be the time step at which the sequence $x_0, Ax_0, A^2x_0, \dots$ approaches a multiple of e_d (the dominant eigenvector of A)

- Now instead of a stochastic matrix let us consider any square matrix A
- Let p be the time step at which the sequence $x_0, Ax_0, A^2x_0, \ldots$ approaches a multiple of e_d (the dominant eigenvector of A)

$$A^p x_0 = k e_d$$

- Now instead of a stochastic matrix let us consider any square matrix A
- Let p be the time step at which the sequence $x_0, Ax_0, A^2x_0, \ldots$ approaches a multiple of e_d (the dominant eigenvector of A)

$$A^p x_0 = k e_d$$
$$A^{p+1} x_0 = A(A^p x_0) = kAe_d = k\lambda_d e_d$$

- Now instead of a stochastic matrix let us consider any square matrix A
- Let p be the time step at which the sequence $x_0, Ax_0, A^2x_0, \ldots$ approaches a multiple of e_d (the dominant eigenvector of A)

$$A^p x_0 = k e_d$$
$$A^{p+1} x_0 = A(A^p x_0) = kAe_d = k\lambda_d e_d$$
$$A^{p+2} x_0 = A(A^{p+1} x_0) = k\lambda_d A e_d = k\lambda_d^2 e_d$$

- Now instead of a stochastic matrix let us consider any square matrix A
- Let p be the time step at which the sequence $x_0, Ax_0, A^2x_0, \ldots$ approaches a multiple of e_d (the dominant eigenvector of A)

$$
\begin{aligned}
A^p x_0 &= k e_d \\
A^{p+1} x_0 &= A(A^p x_0) = k A e_d = k \lambda_d e_d \\
A^{p+2} x_0 &= A(A^{p+1} x_0) = k \lambda_d A e_d = k \lambda_d^2 e_d \\
A^{p+n} x_0 &=
\end{aligned}
$$

- Now instead of a stochastic matrix let us consider any square matrix A
- Let p be the time step at which the sequence $x_0, Ax_0, A^2x_0, \ldots$ approaches a multiple of e_d (the dominant eigenvector of A)

$$
\begin{aligned}
A^p x_0 &= k e_d \\
A^{p+1} x_0 &= A(A^p x_0) = kAe_d = k\lambda_d e_d \\
A^{p+2} x_0 &= A(A^{p+1} x_0) = k\lambda_d A e_d = k\lambda_d^2 e_d \\
A^{p+n} x_0 &= k(\lambda_d)^n e_d
\end{aligned}
$$

- Now instead of a stochastic matrix let us consider any square matrix A
- Let p be the time step at which the sequence $x_0, Ax_0, A^2x_0, \ldots$ approaches a multiple of e_d (the dominant eigenvector of A)

$$
\begin{aligned}
A^p x_0 &= k e_d \\
A^{p+1} x_0 &= A(A^p x_0) = kAe_d = k\lambda_d e_d \\
A^{p+2} x_0 &= A(A^{p+1} x_0) = k\lambda_d A e_d = k\lambda_d^2 e_d \\
A^{p+n} x_0 &= k(\lambda_d)^n e_d
\end{aligned}
$$

- In general, if λ_d is the dominant eigenvalue of a matrix A, what would happen to the sequence $x_0, Ax_0, A^2x_0, \ldots$ if

- Now instead of a stochastic matrix let us consider any square matrix A
- Let p be the time step at which the sequence $x_0, Ax_0, A^2x_0, \ldots$ approaches a multiple of e_d (the dominant eigenvector of A)

$$
\begin{aligned}
A^p x_0 &= k e_d \\
A^{p+1} x_0 &= A(A^p x_0) = kAe_d = k\lambda_d e_d \\
A^{p+2} x_0 &= A(A^{p+1} x_0) = k\lambda_d A e_d = k\lambda_d^2 e_d \\
A^{p+n} x_0 &= k(\lambda_d)^n e_d
\end{aligned}
$$

- In general, if λ_d is the dominant eigenvalue of a matrix A, what would happen to the sequence $x_0, Ax_0, A^2x_0, \ldots$ if
  - $|\lambda_d| > 1$

- Now instead of a stochastic matrix let us consider any square matrix A
- Let p be the time step at which the sequence $x_0, Ax_0, A^2x_0, \ldots$ approaches a multiple of e_d (the dominant eigenvector of A)

$$
\begin{aligned}
A^p x_0 &= k e_d \\
A^{p+1} x_0 &= A(A^p x_0) = kAe_d = k\lambda_d e_d \\
A^{p+2} x_0 &= A(A^{p+1} x_0) = k\lambda_d A e_d = k\lambda_d^2 e_d \\
A^{p+n} x_0 &= k(\lambda_d)^n e_d
\end{aligned}
$$

- In general, if λ_d is the dominant eigenvalue of a matrix A, what would happen to the sequence $x_0, Ax_0, A^2x_0, \ldots$ if
  - $|\lambda_d| > 1$ (will explode)

- Now instead of a stochastic matrix let us consider any square matrix A
- Let p be the time step at which the sequence $x_0, Ax_0, A^2x_0, \ldots$ approaches a multiple of e_d (the dominant eigenvector of A)

$$
\begin{aligned}
A^p x_0 &= k e_d \\
A^{p+1} x_0 &= A(A^p x_0) = kAe_d = k\lambda_d e_d \\
A^{p+2} x_0 &= A(A^{p+1} x_0) = k\lambda_d A e_d = k\lambda_d^2 e_d \\
A^{p+n} x_0 &= k(\lambda_d)^n e_d
\end{aligned}
$$

- In general, if λ_d is the dominant eigenvalue of a matrix A, what would happen to the sequence $x_0, Ax_0, A^2x_0, \ldots$ if
  - $|\lambda_d| > 1$ (will explode)
  - $|\lambda_d| < 1$

- Now instead of a stochastic matrix let us consider any square matrix A
- Let p be the time step at which the sequence $x_0, Ax_0, A^2x_0, \ldots$ approaches a multiple of e_d (the dominant eigenvector of A)

$$\begin{aligned}
A^p x_0 &= k e_d \\
A^{p+1} x_0 &= A(A^p x_0) = kAe_d = k\lambda_d e_d \\
A^{p+2} x_0 &= A(A^{p+1} x_0) = k\lambda_d A e_d = k\lambda_d^2 e_d \\
A^{p+n} x_0 &= k(\lambda_d)^n e_d
\end{aligned}$$

- In general, if λ_d is the dominant eigenvalue of a matrix A, what would happen to the sequence $x_0, Ax_0, A^2x_0, \ldots$ if
  - $|\lambda_d| > 1$ (will explode)
  - $|\lambda_d| < 1$ (will vanish)

- Now instead of a stochastic matrix let us consider any square matrix A
- Let p be the time step at which the sequence $x_0, Ax_0, A^2x_0, \ldots$ approaches a multiple of e_d (the dominant eigenvector of A)

$$
\begin{aligned}
A^p x_0 &= k e_d \\
A^{p+1} x_0 &= A(A^p x_0) = kAe_d = k\lambda_d e_d \\
A^{p+2} x_0 &= A(A^{p+1} x_0) = k\lambda_d A e_d = k\lambda_d^2 e_d \\
A^{p+n} x_0 &= k(\lambda_d)^n e_d
\end{aligned}
$$

- In general, if λ_d is the dominant eigenvalue of a matrix A, what would happen to the sequence $x_0, Ax_0, A^2x_0, \ldots$ if
  - $|\lambda_d| > 1$ (will explode)
  - $|\lambda_d| < 1$ (will vanish)
  - $|\lambda_d| = 1$

- Now instead of a stochastic matrix let us consider any square matrix A
- Let p be the time step at which the sequence $x_0, Ax_0, A^2x_0, \ldots$ approaches a multiple of e_d (the dominant eigenvector of A)

$$
\begin{aligned}
A^p x_0 &= k e_d \\
A^{p+1} x_0 &= A(A^p x_0) = kAe_d = k\lambda_d e_d \\
A^{p+2} x_0 &= A(A^{p+1} x_0) = k\lambda_d A e_d = k\lambda_d^2 e_d \\
A^{p+n} x_0 &= k(\lambda_d)^n e_d
\end{aligned}
$$

- In general, if λ_d is the dominant eigenvalue of a matrix A, what would happen to the sequence $x_0, Ax_0, A^2x_0, \ldots$ if
  - $|\lambda_d| > 1$ (will explode)
  - $|\lambda_d| < 1$ (will vanish)
  - $|\lambda_d| = 1$ (will reach a steady state)

- Now instead of a stochastic matrix let us consider any square matrix A
- Let p be the time step at which the sequence $x_0, Ax_0, A^2x_0, \ldots$ approaches a multiple of e_d (the dominant eigenvector of A)

$$\begin{aligned}
A^p x_0 &= k e_d \\
A^{p+1} x_0 &= A(A^p x_0) = kAe_d = k\lambda_d e_d \\
A^{p+2} x_0 &= A(A^{p+1} x_0) = k\lambda_d A e_d = k\lambda_d^2 e_d \\
A^{p+n} x_0 &= k(\lambda_d)^n e_d
\end{aligned}$$

- In general, if λ_d is the dominant eigenvalue of a matrix A, what would happen to the sequence $x_0, Ax_0, A^2x_0, \ldots$ if
  - $|\lambda_d| > 1$ (will explode)
  - $|\lambda_d| < 1$ (will vanish)
  - $|\lambda_d| = 1$ (will reach a steady state)
- (We will use this in the course at some point)