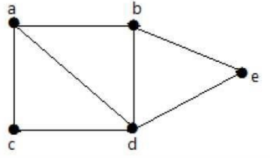
COMP 761: Lecture 12 – Graph Theory II

David Rolnick

September 30, 2020

Problem

In the following graph, I start at vertex *a* and move from where I am to an adjacent vertex each move. What is the probability of being back at *a* after 20 moves?



(Please don't post your ideas in the chat just yet, we'll discuss the problem soon in class.)

Course Announcements

Course Announcements

- Office hours this one Friday moved to 11:30 am-12:30 pm Montreal time.

Finishing up from last class: Eigenvectors

Prove that for any $A \in R^{n \times n}$, eigenvectors with different eigenvalues are linearly independent.

Finishing up from last class: Eigenvectors

Prove that for any $A \in R^{n \times n}$, eigenvectors with different eigenvalues are linearly independent.

- Suppose $v_1, \ldots, v_k$ are some eigenvectors of A with different eigenvalues $\lambda_1, \ldots, \lambda_k$, respectively. What can we do?

Finishing up from last class: Eigenvectors

Prove that for any $A \in R^{n \times n}$, eigenvectors with different eigenvalues are linearly independent.

- Suppose $v_1, \ldots, v_k$ are some eigenvectors of A with different eigenvalues $\lambda_1, \ldots, \lambda_k$, respectively. What can we do?
- Suppose towards contradiction there is a linear dependence:

$$0 = c_1 v_1 + c_2 v_2 + \cdots + c_k v_k.$$

Finishing up from last class: Eigenvectors

Prove that for any $A \in R^{n \times n}$, eigenvectors with different eigenvalues are linearly independent.

- Suppose $v_1, \ldots, v_k$ are some eigenvectors of A with different eigenvalues $\lambda_1, \ldots, \lambda_k$, respectively. What can we do?
- Suppose towards contradiction there is a linear dependence:

$$0 = c_1 v_1 + c_2 v_2 + \cdots + c_k v_k.$$

- Let's multiply both sides by A:

$$0 = c_1 A v_1 + c_2 A v_2 + \cdots + c_k A v_k = c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \cdots + c_k \lambda_k v_k.$$

Finishing up from last class: Eigenvectors

Prove that for any $A \in R^{n \times n}$, eigenvectors with different eigenvalues are linearly independent.

- Suppose $v_1, \ldots, v_k$ are some eigenvectors of A with different eigenvalues $\lambda_1, \ldots, \lambda_k$, respectively. What can we do?
- Suppose towards contradiction there is a linear dependence:

$$0 = c_1 v_1 + c_2 v_2 + \cdots + c_k v_k.$$

- Let's multiply both sides by A:

$$0 = c_1 A v_1 + c_2 A v_2 + \cdots + c_k A v_k = c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \cdots + c_k \lambda_k v_k.$$

- What now?

Finishing up from last class: Eigenvectors

Prove that for any $A \in R^{n \times n}$, eigenvectors with different eigenvalues are linearly independent.

- Suppose $v_1, \ldots, v_k$ are some eigenvectors of A with different eigenvalues $\lambda_1, \ldots, \lambda_k$, respectively. What can we do?
- Suppose towards contradiction there is a linear dependence:

$$0 = c_1 v_1 + c_2 v_2 + \cdots + c_k v_k.$$

- Let's multiply both sides by A:

$$0 = c_1 A v_1 + c_2 A v_2 + \cdots + c_k A v_k = c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \cdots + c_k \lambda_k v_k.$$

- What now?
- Then, we can cancel:

$$\begin{aligned} 0 &= \lambda_k (c_1 v_1 + c_2 v_2 + \cdots + c_k v_k) - (c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \cdots + c_k \lambda_k v_k) \\ &= (\lambda_k c_1 - \lambda_1 c_1) v_1 + (\lambda_k c_2 - \lambda_2 c_2) v_2 + \cdots + (\lambda_k c_{k-1} - \lambda_{k-1} c_{k-1}) v_{k-1} \end{aligned}$$

Finishing up from last class: Eigenvectors

Prove that for any $A \in R^{n \times n}$, eigenvectors with different eigenvalues are linearly independent.

- Suppose $v_1, \ldots, v_k$ are some eigenvectors of A with different eigenvalues $\lambda_1, \ldots, \lambda_k$, respectively. What can we do?
- Suppose towards contradiction there is a linear dependence:

$$0 = c_1 v_1 + c_2 v_2 + \cdots + c_k v_k.$$

- Let's multiply both sides by A:

$$0 = c_1 A v_1 + c_2 A v_2 + \cdots + c_k A v_k = c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \cdots + c_k \lambda_k v_k.$$

- What now?
- Then, we can cancel:

$$\begin{aligned} 0 &= \lambda_k (c_1 v_1 + c_2 v_2 + \cdots + c_k v_k) - (c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \cdots + c_k \lambda_k v_k) \\ &= (\lambda_k c_1 - \lambda_1 c_1) v_1 + (\lambda_k c_2 - \lambda_2 c_2) v_2 + \cdots + (\lambda_k c_{k-1} - \lambda_{k-1} c_{k-1}) v_{k-1} \end{aligned}$$

- We can keep on doing this to reduce the number k.

Finishing up from last class: Eigenvectors

Prove that for any $A \in R^{n \times n}$, eigenvectors with different eigenvalues are linearly independent.

- Suppose $v_1, \ldots, v_k$ are some eigenvectors of A with different eigenvalues $\lambda_1, \ldots, \lambda_k$, respectively. What can we do?
- Suppose towards contradiction there is a linear dependence:

$$0 = c_1 v_1 + c_2 v_2 + \cdots + c_k v_k.$$

- Let's multiply both sides by A:

$$0 = c_1 A v_1 + c_2 A v_2 + \cdots + c_k A v_k = c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \cdots + c_k \lambda_k v_k.$$

- What now?
- Then, we can cancel:

$$\begin{aligned} 0 &= \lambda_k (c_1 v_1 + c_2 v_2 + \cdots + c_k v_k) - (c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \cdots + c_k \lambda_k v_k) \\ &= (\lambda_k c_1 - \lambda_1 c_1) v_1 + (\lambda_k c_2 - \lambda_2 c_2) v_2 + \cdots + (\lambda_k c_{k-1} - \lambda_{k-1} c_{k-1}) v_{k-1} \end{aligned}$$

- We can keep on doing this to reduce the number k.
- Clean way to write proof - extremal principle, suppose we picked k minimal such that there is a linear dependence.

Symmetric matrices

Prove that if $A \in R^{n \times n}$ is symmetric (that is, $A = A^T$), then eigenvectors with different eigenvalues are orthogonal.

Symmetric matrices

Prove that if $A \in R^{n \times n}$ is symmetric (that is, $A = A^T$), then eigenvectors with different eigenvalues are orthogonal.

- Suppose v_1, v_2 are eigenvectors of A with eigenvalues λ_1, λ_2, respectively.

Symmetric matrices

Prove that if $A \in R^{n \times n}$ is symmetric (that is, $A = A^T$), then eigenvectors with different eigenvalues are orthogonal.

- Suppose v_1, v_2 are eigenvectors of A with eigenvalues λ_1, λ_2, respectively.
- We want to show $0 = v_1 \cdot v_2 = v_1^T v_2$.

Symmetric matrices

Prove that if $A \in R^{n \times n}$ is symmetric (that is, $A = A^T$), then eigenvectors with different eigenvalues are orthogonal.

- Suppose v_1, v_2 are eigenvectors of A with eigenvalues λ_1, λ_2, respectively.
- We want to show $0 = v_1 \cdot v_2 = v_1^T v_2$.
- Since v_1 and v_2 are eigenvectors:

$$\begin{aligned} \lambda_1 v_1^T v_2 &= (Av_1)^T v_2 \\ &= v_1^T A^T v_2 \\ &= v_1^T A v_2 \text{ since } A \text{ is symmetric} \\ &= v_1^T (\lambda_2 v_2). \end{aligned}$$

Symmetric matrices

Prove that if $A \in R^{n \times n}$ is symmetric (that is, $A = A^T$), then eigenvectors with different eigenvalues are orthogonal.

- Suppose v_1, v_2 are eigenvectors of A with eigenvalues λ_1, λ_2, respectively.
- We want to show $0 = v_1 \cdot v_2 = v_1^T v_2$.
- Since v_1 and v_2 are eigenvectors:

$$\begin{aligned}
\lambda_1 v_1^T v_2 &= (Av_1)^T v_2 \\
&= v_1^T A^T v_2 \\
&= v_1^T A v_2 \text{ since } A \text{ is symmetric} \\
&= v_1^T (\lambda_2 v_2).
\end{aligned}$$

- So $(\lambda_1 - \lambda_2) v_1^T v_2 = 0$.

Symmetric matrices

Prove that if $A \in R^{n\times n}$ is symmetric (that is, $A = A^T$), then eigenvectors with different eigenvalues are orthogonal.

- Suppose v_1, v_2 are eigenvectors of A with eigenvalues λ_1, λ_2, respectively.
- We want to show $0 = v_1 \cdot v_2 = v_1^T v_2$.
- Since v_1 and v_2 are eigenvectors:

$$\begin{aligned} \lambda_1 v_1^T v_2 &= (Av_1)^T v_2 \\ &= v_1^T A^T v_2 \\ &= v_1^T A v_2 \text{ since } A \text{ is symmetric} \\ &= v_1^T(\lambda_2 v_2). \end{aligned}$$

- So $(\lambda_1 - \lambda_2)v_1^T v_2 = 0$.
- Since $\lambda_1 \neq \lambda_2$, we conclude $v_1^T v_2 = 0$.

Symmetric matrices

Prove that if $A \in R^{n \times n}$ is symmetric (that is, $A = A^T$), then eigenvectors with different eigenvalues are orthogonal.

- Suppose v_1, v_2 are eigenvectors of A with eigenvalues λ_1, λ_2, respectively.
- We want to show $0 = v_1 \cdot v_2 = v_1^T v_2$.
- Since v_1 and v_2 are eigenvectors:

$$\begin{aligned} \lambda_1 v_1^T v_2 &= (Av_1)^T v_2 \\ &= v_1^T A^T v_2 \\ &= v_1^T A v_2 \text{ since } A \text{ is symmetric} \\ &= v_1^T (\lambda_2 v_2). \end{aligned}$$

- So $(\lambda_1 - \lambda_2) v_1^T v_2 = 0$.
- Since $\lambda_1 \neq \lambda_2$, we conclude $v_1^T v_2 = 0$.
- So v_1 and v_2 are orthogonal.

Symmetric matrices

Prove that if a matrix $A \in \mathbb{R}^{n \times n}$ is symmetric, then all its eigenvalues are real.

Symmetric matrices

Prove that if a matrix $A \in \mathbb{R}^{n \times n}$ is symmetric, then all its eigenvalues are real.

- Suppose v, λ are an eigenvector/eigenvalue pair of A.

Symmetric matrices

Prove that if a matrix $A \in \mathbb{R}^{n \times n}$ is symmetric, then all its eigenvalues are real.

- Suppose v, λ are an eigenvector/eigenvalue pair of A.
- The *complex conjugate* of a complex number $x = a + bi$ is the number $\bar{x} = a - bi$.

Symmetric matrices

Prove that if a matrix $A \in \mathbb{R}^{n\times n}$ is symmetric, then all its eigenvalues are real.

- Suppose v, λ are an eigenvector/eigenvalue pair of A.
- The *complex conjugate* of a complex number $x = a + bi$ is the number $\bar{x} = a - bi$.
- We saw that if a complex number λ is a root of a real polynomial, its conjugate $\bar{\lambda}$ is a root too.

Symmetric matrices

Prove that if a matrix $A \in \mathbb{R}^{n \times n}$ is symmetric, then all its eigenvalues are real.

- Suppose v, λ are an eigenvector/eigenvalue pair of A.
- The *complex conjugate* of a complex number $x = a + bi$ is the number $\bar{x} = a - bi$.
- We saw that if a complex number λ is a root of a real polynomial, its conjugate $\bar{\lambda}$ is a root too.
- Can show using that fact that $\bar{v}, \bar{\lambda}$ are an eigenvector/eigenvalue pair.

Symmetric matrices

Prove that if a matrix $A \in \mathbb{R}^{n \times n}$ is symmetric, then all its eigenvalues are real.

- Suppose v, λ are an eigenvector/eigenvalue pair of A.
- The *complex conjugate* of a complex number $x = a + bi$ is the number $\bar{x} = a - bi$.
- We saw that if a complex number λ is a root of a real polynomial, its conjugate $\bar{\lambda}$ is a root too.
- Can show using that fact that $\bar{v}, \bar{\lambda}$ are an eigenvector/eigenvalue pair.
- If $\bar{\lambda} \neq \lambda$, then $\bar{v}, v$ orthogonal, so:

$$\bar{v}^T v = 0.$$

Symmetric matrices

Prove that if a matrix $A \in \mathbb{R}^{n \times n}$ is symmetric, then all its eigenvalues are real.

- Suppose v, λ are an eigenvector/eigenvalue pair of A.
- The *complex conjugate* of a complex number $x = a + bi$ is the number $\bar{x} = a - bi$.
- We saw that if a complex number λ is a root of a real polynomial, its conjugate $\bar{\lambda}$ is a root too.
- Can show using that fact that $\bar{v}, \bar{\lambda}$ are an eigenvector/eigenvalue pair.
- If $\bar{\lambda} \neq \lambda$, then $\bar{v}, v$ orthogonal, so:

$$\bar{v}^T v = 0.$$

- Not possible, since $(a + bi)(a - bi) = a^2 + b^2 \geq 0$.

Symmetric matrices

Prove that if a matrix $A \in \mathbb{R}^{n \times n}$ is symmetric, then all its eigenvalues are real.

- Suppose v, λ are an eigenvector/eigenvalue pair of A.
- The *complex conjugate* of a complex number $x = a + bi$ is the number $\bar{x} = a - bi$.
- We saw that if a complex number λ is a root of a real polynomial, its conjugate $\bar{\lambda}$ is a root too.
- Can show using that fact that $\bar{v}, \bar{\lambda}$ are an eigenvector/eigenvalue pair.
- If $\bar{\lambda} \neq \lambda$, then $\bar{v}, v$ orthogonal, so:

$$\bar{v}^T v = 0.$$

- Not possible, since $(a + bi)(a - bi) = a^2 + b^2 \geq 0$.
- Therefore $\bar{\lambda} = \lambda$, so λ is actually real.

Powers of a matrix

We write A^k for A multiplied by itself k times. Suppose that $A \in \mathbb{R}^{n \times n}$ has all different eigenvalues. What is $A^k v$ for a general vector v for large k approaching infinity?

Powers of a matrix

We write A^k for A multiplied by itself k times. Suppose that $A \in \mathbb{R}^{n \times n}$ has all different eigenvalues. What is $A^k v$ for a general vector v for large k approaching infinity?

- We know A has n linearly independent eigenvectors $v_1, \ldots, v_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, respectively.

Powers of a matrix

We write A^k for A multiplied by itself k times. Suppose that $A \in \mathbb{R}^{n \times n}$ has all different eigenvalues. What is $A^k v$ for a general vector v for large k approaching infinity?

- We know A has n linearly independent eigenvectors $v_1, \ldots, v_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, respectively.
- Since $Av_j = \lambda_j v_j$, we have $A^2 v_j = \lambda_j^2 v_j$ and $A^3 v_j = \lambda_j^3 v_j$, etc.

Powers of a matrix

We write A^k for A multiplied by itself k times. Suppose that $A \in \mathbb{R}^{n\times n}$ has all different eigenvalues. What is $A^k v$ for a general vector v for large k approaching infinity?

- We know A has n linearly independent eigenvectors $v_1, \ldots, v_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, respectively.
- Since $Av_j = \lambda_j v_j$, we have $A^2 v_j = \lambda_j^2 v_j$ and $A^3 v_j = \lambda_j^3 v_j$, etc.
- Each v_j must be an eigenvector of A^k with eigenvalue λ_j^k.

Powers of a matrix

We write A^k for A multiplied by itself k times. Suppose that $A \in \mathbb{R}^{n \times n}$ has all different eigenvalues. What is $A^k v$ for a general vector v for large k approaching infinity?

- We know A has n linearly independent eigenvectors $v_1, \ldots, v_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, respectively.
- Since $Av_j = \lambda_j v_j$, we have $A^2 v_j = \lambda_j^2 v_j$ and $A^3 v_j = \lambda_j^3 v_j$, etc.
- Each v_j must be an eigenvector of A^k with eigenvalue λ_j^k.
- Let's write v as a linear combination of the eigenvectors:

$$v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n.$$

Powers of a matrix

We write A^k for A multiplied by itself k times. Suppose that $A \in \mathbb{R}^{n \times n}$ has all different eigenvalues. What is $A^k v$ for a general vector v for large k approaching infinity?

- We know A has n linearly independent eigenvectors $v_1, \ldots, v_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, respectively.
- Since $Av_j = \lambda_j v_j$, we have $A^2 v_j = \lambda_j^2 v_j$ and $A^3 v_j = \lambda_j^3 v_j$, etc.
- Each v_j must be an eigenvector of A^k with eigenvalue λ_j^k.
- Let's write v as a linear combination of the eigenvectors:

$$v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n.$$

- Then, multiplying by A^k, we have:

$$A^k v = c_1 \lambda_1^k v_1 + c_2 \lambda_2^k v_2 + \cdots + c_n \lambda_n^k v_n.$$

Powers of a matrix

We write A^k for A multiplied by itself k times. Suppose that $A \in \mathbb{R}^{n \times n}$ has all different eigenvalues. What is $A^k v$ for a general vector v for large k approaching infinity?

- We know A has n linearly independent eigenvectors $v_1, \ldots, v_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, respectively.
- Since $Av_j = \lambda_j v_j$, we have $A^2 v_j = \lambda_j^2 v_j$ and $A^3 v_j = \lambda_j^3 v_j$, etc.
- Each v_j must be an eigenvector of A^k with eigenvalue λ_j^k.
- Let's write v as a linear combination of the eigenvectors:

$$v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n.$$

- Then, multiplying by A^k, we have:

$$A^k v = c_1 \lambda_1^k v_1 + c_2 \lambda_2^k v_2 + \cdots + c_n \lambda_n^k v_n.$$

- For k big, the term that dominates is the one where $|\lambda_j|$ is largest.

Powers of a matrix

We write A^k for A multiplied by itself k times. Suppose that $A \in \mathbb{R}^{n \times n}$ has all different eigenvalues. What is $A^k v$ for a general vector v for large k approaching infinity?

- We know A has n linearly independent eigenvectors $v_1, \ldots, v_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, respectively.
- Since $Av_j = \lambda_j v_j$, we have $A^2 v_j = \lambda_j^2 v_j$ and $A^3 v_j = \lambda_j^3 v_j$, etc.
- Each v_j must be an eigenvector of A^k with eigenvalue λ_j^k.
- Let's write v as a linear combination of the eigenvectors:

$$v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n.$$

- Then, multiplying by A^k, we have:

$$A^k v = c_1 \lambda_1^k v_1 + c_2 \lambda_2^k v_2 + \cdots + c_n \lambda_n^k v_n.$$

- For k big, the term that dominates is the one where $|\lambda_j|$ is largest.
- So $A^k v$ approaches a multiple of the *top eigenvector* v_j (the one where $|\lambda_j|$ is largest).

Principal Component Analysis

Principal Component Analysis

- Sometimes there is very high-dimensional data and we would like to approximate it using a lower number of dimensions.

Principal Component Analysis

- Sometimes there is very high-dimensional data and we would like to approximate it using a lower number of dimensions.
- *Principal component analysis* (PCA) is one way to do this.

Principal Component Analysis

- Sometimes there is very high-dimensional data and we would like to approximate it using a lower number of dimensions.
- *Principal component analysis* (PCA) is one way to do this.
- Suppose the datapoints are $x_1, x_2, \ldots, x_m$, each a vector in $\mathbb{R}^n$.

Principal Component Analysis

- Sometimes there is very high-dimensional data and we would like to approximate it using a lower number of dimensions.
- *Principal component analysis* (PCA) is one way to do this.
- Suppose the datapoints are $x_1, x_2, \ldots, x_m$, each a vector in $\mathbb{R}^n$.
- Represent the data as a matrix $X \in \mathbb{R}^{m \times n}$.

Principal Component Analysis

- Sometimes there is very high-dimensional data and we would like to approximate it using a lower number of dimensions.
- *Principal component analysis* (PCA) is one way to do this.
- Suppose the datapoints are $x_1, x_2, \ldots, x_m$, each a vector in $\mathbb{R}^n$.
- Represent the data as a matrix $X \in \mathbb{R}^{m \times n}$.
- *Covariance matrix* is $X^T X \in \mathbb{R}^{n \times n}$ (pairwise similarities in data).

Principal Component Analysis

- Sometimes there is very high-dimensional data and we would like to approximate it using a lower number of dimensions.
- *Principal component analysis* (PCA) is one way to do this.
- Suppose the datapoints are $x_1, x_2, \ldots, x_m$, each a vector in $\mathbb{R}^n$.
- Represent the data as a matrix $X \in \mathbb{R}^{m \times n}$.
- *Covariance matrix* is $X^T X \in \mathbb{R}^{n \times n}$ (pairwise similarities in data).
- Since it is symmetric, all real eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$.

Principal Component Analysis

- Sometimes there is very high-dimensional data and we would like to approximate it using a lower number of dimensions.
- *Principal component analysis* (PCA) is one way to do this.
- Suppose the datapoints are $x_1, x_2, \ldots, x_m$, each a vector in $\mathbb{R}^n$.
- Represent the data as a matrix $X \in \mathbb{R}^{m \times n}$.
- *Covariance matrix* is $X^T X \in \mathbb{R}^{n \times n}$ (pairwise similarities in data).
- Since it is symmetric, all real eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$.
- Eigenvectors $v_1, v_2, \ldots, v_n$, e.g. v_1 is the direction that varies most across datapoints. These are called *principal components*.
- Express each x_k as:

$$x_k = c_{k1} v_1 + c_{k2} v_2 + \cdots + c_{kn} v_n.$$

Principal Component Analysis

- Sometimes there is very high-dimensional data and we would like to approximate it using a lower number of dimensions.
- *Principal component analysis* (PCA) is one way to do this.
- Suppose the datapoints are $x_1, x_2, \ldots, x_m$, each a vector in $\mathbb{R}^n$.
- Represent the data as a matrix $X \in \mathbb{R}^{m \times n}$.
- *Covariance matrix* is $X^T X \in \mathbb{R}^{n \times n}$ (pairwise similarities in data).
- Since it is symmetric, all real eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$.
- Eigenvectors $v_1, v_2, \ldots, v_n$, e.g. v_1 is the direction that varies most across datapoints. These are called *principal components*.
- Express each x_k as:

$$x_k = c_{k1} v_1 + c_{k2} v_2 + \cdots + c_{kn} v_n.$$

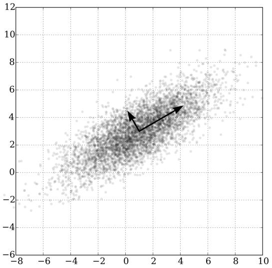
- Lot of variation captured in just first few coefficients $c_{k1}, c_{k2}, \ldots$.

Principal Component Analysis

- *Covariance matrix* is $X^T X \in \mathbb{R}^{n\times n}$ (pairwise similarities in data).
- Since it is symmetric, all real eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$.
- Eigenvectors $v_1, v_2, \ldots, v_n$, e.g. v_1 is the direction that varies most across datapoints. These are called *principal components*.
- Express each x_k as:

$$x_k = c_{k1}v_1 + c_{k2}v_2 + \cdots + c_{kn}v_n.$$

- Lot of variation captured in just first few coefficients $c_{k1}, c_{k2}, \ldots$.



The adjacency matrix

The adjacency matrix

- Let G be a graph with n vertices.

The adjacency matrix

- Let G be a graph with n vertices.
- The *adjacency matrix* $\text{adj}(G)$ is an $n \times n$ matrix $[a_{ij}]$ where

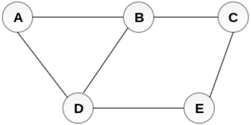
$$a_{ij} = \begin{cases} 1 & \text{if } (i,j) \text{ is an edge} \\ 0 & \text{otherwise} \end{cases}$$

The adjacency matrix

- Let G be a graph with n vertices.
- The *adjacency matrix* $\mathrm{adj}(G)$ is an $n \times n$ matrix $[a_{ij}]$ where

$$a_{ij} = \begin{cases} 1 & \text{if } (i,j) \text{ is an edge} \\ 0 & \text{otherwise} \end{cases}$$

- Example:



Undirected Graph

	A	B	C	D	E
A	0	1	0	1	0
B	1	0	1	1	0
C	0	1	0	0	1
D	1	1	0	0	1
E	0	0	1	1	0

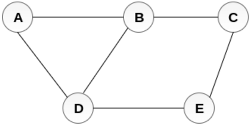
Adjacency Matrix

The adjacency matrix

- Let G be a graph with n vertices.
- The *adjacency matrix* $\text{adj}(G)$ is an $n \times n$ matrix $[a_{ij}]$ where

$$a_{ij} = \begin{cases} 1 & \text{if } (i,j) \text{ is an edge} \\ 0 & \text{otherwise} \end{cases}$$

- Example:



Undirected Graph

	A	B	C	D	E
A	0	1	0	1	0
B	1	0	1	1	0
C	0	1	0	0	1
D	1	1	0	0	1
E	0	0	1	1	0

Adjacency Matrix

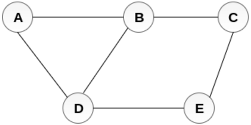
- The adjacency matrix is symmetric (for undirected graphs).

The adjacency matrix

- Let G be a graph with n vertices.
- The *adjacency matrix* $\text{adj}(G)$ is an $n \times n$ matrix $[a_{ij}]$ where

$$a_{ij} = \begin{cases} 1 & \text{if } (i, j) \text{ is an edge} \\ 0 & \text{otherwise} \end{cases}$$

- Example:



Undirected Graph

	A	B	C	D	E
A	0	1	0	1	0
B	1	0	1	1	0
C	0	1	0	0	1
D	1	1	0	0	1
E	0	0	1	1	0

Adjacency Matrix

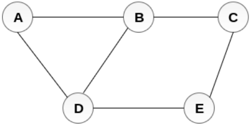
- The adjacency matrix is symmetric (for undirected graphs).
- What happens if we sum all rows of $\text{adj}(G)$?

The adjacency matrix

- Let G be a graph with n vertices.
- The *adjacency matrix* $\text{adj}(G)$ is an $n \times n$ matrix $[a_{ij}]$ where

$$a_{ij} = \begin{cases} 1 & \text{if } (i,j) \text{ is an edge} \\ 0 & \text{otherwise} \end{cases}$$

- Example:



Undirected Graph

	A	B	C	D	E
A	0	1	0	1	0
B	1	0	1	1	0
C	0	1	0	0	1
D	1	1	0	0	1
E	0	0	1	1	0

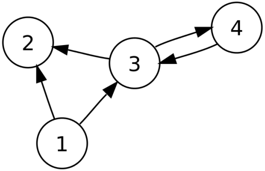
Adjacency Matrix

- The adjacency matrix is symmetric (for undirected graphs).
- What happens if we sum all rows of $\text{adj}(G)$?
- We have the vector of degrees of vertices.

Directed graphs

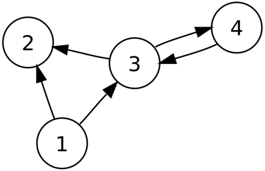
Directed graphs

- In *directed graphs*, each edge in an *ordered* pair – e.g. $(1, 2)$ is not the same as $(2, 1)$. (Can think of edges having a *direction*.)



Directed graphs

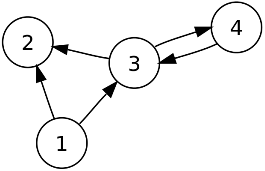
- In *directed graphs*, each edge in an *ordered* pair – e.g. $(1, 2)$ is not the same as $(2, 1)$. (Can think of edges having a *direction*.)



- Examples: Social networks, movement between places, dependency relationships, etc.

Directed graphs

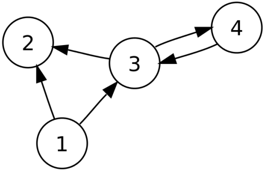
- In *directed graphs*, each edge in an *ordered* pair – e.g. $(1, 2)$ is not the same as $(2, 1)$. (Can think of edges having a *direction*.)



- Examples: Social networks, movement between places, dependency relationships, etc.
- What happens to the adjacency matrix?

Directed graphs

- In *directed graphs*, each edge in an *ordered* pair – e.g. $(1, 2)$ is not the same as $(2, 1)$. (Can think of edges having a *direction*.)



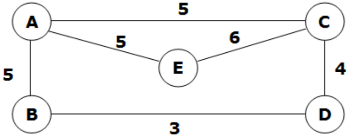
- Examples: Social networks, movement between places, dependency relationships, etc.
- What happens to the adjacency matrix?
- No longer symmetric:

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

Weighted graphs

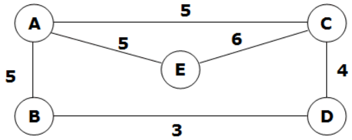
Weighted graphs

- In *weighted graphs*, each edge has a *weight* associated to it.



Weighted graphs

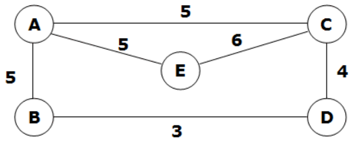
- In *weighted graphs*, each edge has a *weight* associated to it.



- Examples: Lengths of roads, resistances of wires, pairwise similarities, etc.

Weighted graphs

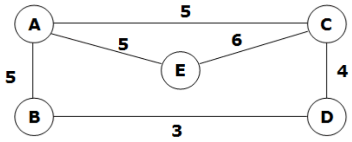
- In *weighted graphs*, each edge has a *weight* associated to it.



- Examples: Lengths of roads, resistances of wires, pairwise similarities, etc.
- What happens to the adjacency matrix?

Weighted graphs

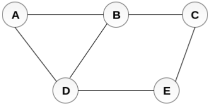
- In *weighted graphs*, each edge has a *weight* associated to it.



- Examples: Lengths of roads, resistances of wires, pairwise similarities, etc.
- What happens to the adjacency matrix?
- No longer binary:

$$\begin{bmatrix} 0 & 5 & 5 & 0 & 5 \\ 5 & 0 & 0 & 3 & 0 \\ 5 & 0 & 0 & 4 & 6 \\ 0 & 3 & 4 & 0 & 0 \\ 5 & 0 & 6 & 0 & 0 \end{bmatrix}$$

Random walks

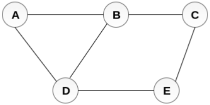


Undirected Graph

	A	B	C	D	E
A	0	1	0	1	0
B	1	0	1	1	0
C	0	1	0	0	1
D	1	1	0	0	1
E	0	0	1	1	0

Adjacency Matrix

Random walks



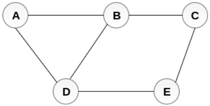
Undirected Graph

	A	B	C	D	E
A	0	1	0	1	0
B	1	0	1	1	0
C	0	1	0	0	1
D	1	1	0	0	1
E	0	0	1	1	0

Adjacency Matrix

- Often, graphs represent flow of information / objects in a network.

Random walks



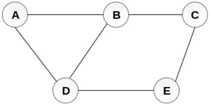
Undirected Graph

	A	B	C	D	E
A	0	1	0	1	0
B	1	0	1	1	0
C	0	1	0	0	1
D	1	1	0	0	1
E	0	0	1	1	0

Adjacency Matrix

- Often, graphs represent flow of information / objects in a network.
- Why is the adjacency matrix relevant to this?

Random walks



Undirected Graph

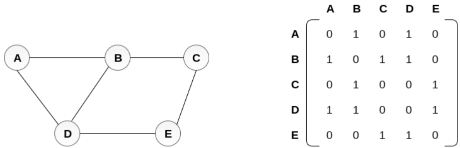
	A	B	C	D	E
A	0	1	0	1	0
B	1	0	1	1	0
C	0	1	0	0	1
D	1	1	0	0	1
E	0	0	1	1	0

Adjacency Matrix

- Often, graphs represent flow of information / objects in a network.
- Why is the adjacency matrix relevant to this?
- Let's look at the square of the adjacency matrix:

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 & 1 & 1 \\ 1 & 3 & 0 & 1 & 2 \\ 1 & 0 & 2 & 2 & 0 \\ 1 & 1 & 2 & 3 & 0 \\ 1 & 2 & 0 & 0 & 2 \end{bmatrix}$$

Random walks



Undirected Graph   Adjacency Matrix

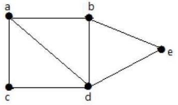
- Often, graphs represent flow of information / objects in a network.
- Why is the adjacency matrix relevant to this?
- Let's look at the square of the adjacency matrix:

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 & 1 & 1 \\ 1 & 3 & 0 & 1 & 2 \\ 1 & 0 & 2 & 2 & 0 \\ 1 & 1 & 2 & 3 & 0 \\ 1 & 2 & 0 & 0 & 2 \end{bmatrix}$$

- Entry in row *i*, column *j* is $\#$ of 2-edge paths between *i* and *j*.

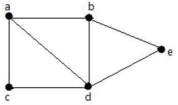
Random walks

In the following graph, I start at vertex *a* and move from where I am to an adjacent vertex each move. What is the probability of being back at *a* after 10 moves?



Random walks

In the following graph, I start at vertex *a* and move from where I am to an adjacent vertex each move. What is the probability of being back at *a* after 10 moves?

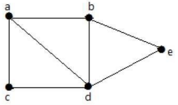


- We know that the *k*th power of the adjacency matrix *A* gives the number of *k*-edge paths between vertices.

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}, \; A^{10} = \begin{bmatrix} 10537 & 10448 & 7801 & 12475 & 7856 \\ 10448 & 10537 & 7856 & 12475 & 7801 \\ 7801 & 7856 & 5863 & 9293 & 5829 \\ 12475 & 12475 & 9293 & 14874 & 9293 \\ 7856 & 7801 & 5829 & 9293 & 5863 \end{bmatrix}$$

Random walks

In the following graph, I start at vertex a and move from where I am to an adjacent vertex each move. What is the probability of being back at a after 10 moves?



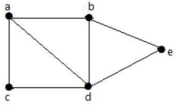
- We know that the kth power of the adjacency matrix A gives the number of k-edge paths between vertices.

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}, \ A^{10} = \begin{bmatrix} 10537 & 10448 & 7801 & 12475 & 7856 \\ 10448 & 10537 & 7856 & 12475 & 7801 \\ 7801 & 7856 & 5863 & 9293 & 5829 \\ 12475 & 12475 & 9293 & 14874 & 9293 \\ 7856 & 7801 & 5829 & 9293 & 5863 \end{bmatrix}$$

- (Faster way to get A^{10}: calculate A^2, A^4, A^8, use $A^{10} = (A^2)(A^8)$.)

Random walks

In the following graph, I start at vertex a and move from where I am to an adjacent vertex each move. What is the probability of being back at a after 10 moves?



- We know that the kth power of the adjacency matrix A gives the number of k-edge paths between vertices.

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}, \; A^{10} = \begin{bmatrix} 10537 & 10448 & 7801 & 12475 & 7856 \\ 10448 & 10537 & 7856 & 12475 & 7801 \\ 7801 & 7856 & 5863 & 9293 & 5829 \\ 12475 & 12475 & 9293 & 14874 & 9293 \\ 7856 & 7801 & 5829 & 9293 & 5863 \end{bmatrix}$$

- (Faster way to get A^{10}: calculate A^2, A^4, A^8, use $A^{10} = (A^2)(A^8)$.)
- Answer: $10537/(10537 + 10448 + 7801 + 12475 + 7856) \approx 0.215$.

Random walks

Random walks

- We just look at the first column of A^{10} – the different entries give the relative likelihood of being at different vertices after 10 steps.

Random walks

- We just look at the first column of A^{10} – the different entries give the relative likelihood of being at different vertices after 10 steps.
- What happens if we take a very large number of random steps on a connected graph, starting at vertex 1?

Random walks

- We just look at the first column of A^{10} – the different entries give the relative likelihood of being at different vertices after 10 steps.
- What happens if we take a very large number of random steps on a connected graph, starting at vertex 1?
- Look at A^k for $k \to \infty$.

Random walks

- We just look at the first column of A^{10} – the different entries give the relative likelihood of being at different vertices after 10 steps.
- What happens if we take a very large number of random steps on a connected graph, starting at vertex 1?
- Look at A^k for $k \to \infty$.
- The first column of A^k equals $A^k v_1$, where

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Random walks

- We just look at the first column of A^{10} – the different entries give the relative likelihood of being at different vertices after 10 steps.
- What happens if we take a very large number of random steps on a connected graph, starting at vertex 1?
- Look at A^k for $k \to \infty$.
- The first column of A^k equals $A^k v_1$, where

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

- We know that $A^k v_1$ will approach cv, where v is the eigenvector with highest eigenvalue.

Random walks

- We just look at the first column of A^{10} – the different entries give the relative likelihood of being at different vertices after 10 steps.
- What happens if we take a very large number of random steps on a connected graph, starting at vertex 1?
- Look at A^k for $k \to \infty$.
- The first column of A^k equals $A^k v_1$, where

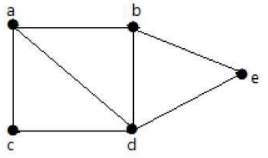
$$v_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

- We know that $A^k v_1$ will approach cv, where v is the eigenvector with highest eigenvalue.
- Therefore a long random walk on the graph will approach a limiting distribution among the vertices, given by the top eigenvector.

Diameter of a graph

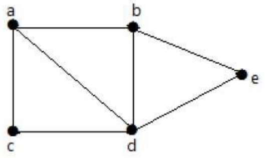
Diameter of a graph

- The *distance* between two vertices in a graph is the length of the shortest path between them.



Diameter of a graph

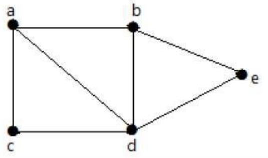
- The *distance* between two vertices in a graph is the length of the shortest path between them.



- Example: Distance between a and e is 2.

Diameter of a graph

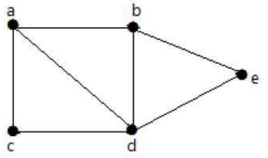
- The *distance* between two vertices in a graph is the length of the shortest path between them.



- Example: Distance between a and e is 2.
- The *diameter* of a graph G is the maximum distance between any two vertices.

Diameter of a graph

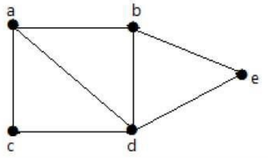
- The *distance* between two vertices in a graph is the length of the shortest path between them.



- Example: Distance between a and e is 2.
- The *diameter* of a graph G is the maximum distance between any two vertices.
- Example: "Six degrees of separation"

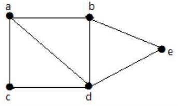
Diameter of a graph

- The *distance* between two vertices in a graph is the length of the shortest path between them.

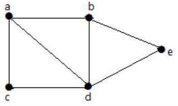


- Example: Distance between a and e is 2.
- The *diameter* of a graph G is the maximum distance between any two vertices.
- Example: "Six degrees of separation"
- Diameter is the lowest k such that $(\text{adj}(G))^k$ has all nonzero entries.

The Laplacian matrix



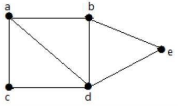
The Laplacian matrix



- The *Laplacian matrix* of a graph G is defined as $D - A$, where A is the adjacency matrix and D is the diagonal matrix of vertex degrees:

$$D = \begin{bmatrix} \deg(v_1) & 0 & \cdots & 0 \\ 0 & \deg(v_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \deg(v_n) \end{bmatrix}.$$

The Laplacian matrix



- The *Laplacian matrix* of a graph G is defined as $D - A$, where A is the adjacency matrix and D is the diagonal matrix of vertex degrees:

$$D = \begin{bmatrix} \deg(v_1) & 0 & \cdots & 0 \\ 0 & \deg(v_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \deg(v_n) \end{bmatrix}.$$

- Example for the graph above:

$$L = \begin{bmatrix} 3 & -1 & -1 & -1 & 0 \\ -1 & 3 & 0 & -1 & -1 \\ -1 & 0 & 2 & -1 & 0 \\ -1 & -1 & -1 & 4 & -1 \\ 0 & -1 & 0 & -1 & 2 \end{bmatrix}.$$

The Laplacian matrix

- The *Laplacian matrix* of a graph G is defined as $D - A$, where A is the adjacency matrix and D is the diagonal matrix of vertex degrees.

The Laplacian matrix

- The *Laplacian matrix* of a graph G is defined as $D - A$, where A is the adjacency matrix and D is the diagonal matrix of vertex degrees.
- What is one eigenvector for the Laplacian?

The Laplacian matrix

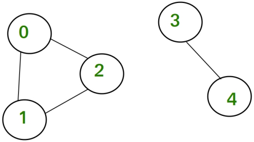
- The *Laplacian matrix* of a graph G is defined as $D - A$, where A is the adjacency matrix and D is the diagonal matrix of vertex degrees.
- What is one eigenvector for the Laplacian?
- The sum of each row is 0, since for every edge, there is a -1 in A and a $+1$ in D.

The Laplacian matrix

- The *Laplacian matrix* of a graph G is defined as $D - A$, where A is the adjacency matrix and D is the diagonal matrix of vertex degrees.
- What is one eigenvector for the Laplacian?
- The sum of each row is 0, since for every edge, there is a -1 in A and a $+1$ in D.
- That means $[1, 1, \ldots, 1]$ is an eigenvector with eigenvalue 0.

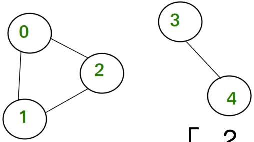
The Laplacian matrix

What if the graph is disconnected?



The Laplacian matrix

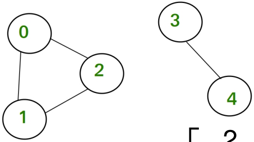
What if the graph is disconnected?



- The Laplacian for this graph is $L = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}.$

The Laplacian matrix

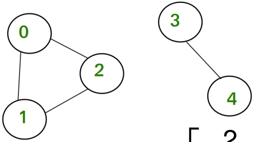
What if the graph is disconnected?



- The Laplacian for this graph is $L = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}$.
- Any other eigenvectors in addition to $[1, 1, 1, 1, 1]$?

The Laplacian matrix

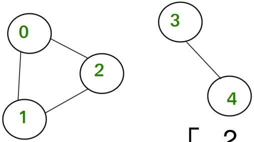
What if the graph is disconnected?



- The Laplacian for this graph is $L = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}$.
- Any other eigenvectors in addition to $[1, 1, 1, 1, 1]$?
- There are also eigenvectors $[1, 1, 1, 0, 0]$ and $[0, 0, 0, 1, 1]$, both with eigenvalue 0.

The Laplacian matrix

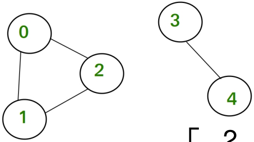
What if the graph is disconnected?



- The Laplacian for this graph is $L = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}$.
- Any other eigenvectors in addition to $[1, 1, 1, 1, 1]$?
- There are also eigenvectors $[1, 1, 1, 0, 0]$ and $[0, 0, 0, 1, 1]$, both with eigenvalue 0.
- They are orthogonal and $[1, 1, 1, 1, 1] = [1, 1, 1, 0, 0] + [0, 0, 0, 1, 1]$

The Laplacian matrix

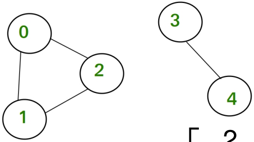
What if the graph is disconnected?



- The Laplacian for this graph is $L = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}$.
- Any other eigenvectors in addition to $[1, 1, 1, 1, 1]$?
- There are also eigenvectors $[1, 1, 1, 0, 0]$ and $[0, 0, 0, 1, 1]$, both with eigenvalue 0.
- They are orthogonal and $[1, 1, 1, 1, 1] = [1, 1, 1, 0, 0] + [0, 0, 0, 1, 1]$
- In general, there is one eigenvector with eigenvalue 0 for each connected component of the graph.

The Laplacian matrix

What if the graph is disconnected?



- The Laplacian for this graph is $L = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}$.
- Any other eigenvectors in addition to $[1, 1, 1, 1, 1]$?
- There are also eigenvectors $[1, 1, 1, 0, 0]$ and $[0, 0, 0, 1, 1]$, both with eigenvalue 0.
- They are orthogonal and $[1, 1, 1, 1, 1] = [1, 1, 1, 0, 0] + [0, 0, 0, 1, 1]$
- In general, there is one eigenvector with eigenvalue 0 for each connected component of the graph.
- Won't prove it, but all eigenvalues ≥ 0 (positive semi-definite), so eigenvalue 0 is smallest.

Spectral clustering

How best to divide a connected graph into two?

Spectral clustering

How best to divide a connected graph into two?

- Suppose the graph is almost disconnected.

Spectral clustering

How best to divide a connected graph into two?

- Suppose the graph is almost disconnected.
- Then there is an eigenvector $[1, 1, \ldots, 1]$ and no other eigenvectors with eigenvalue 0.

Spectral clustering

How best to divide a connected graph into two?

- Suppose the graph is almost disconnected.
- Then there is an eigenvector $[1, 1, \ldots, 1]$ and no other eigenvectors with eigenvalue 0.
- But there is another eigenvector that is very close to $[1, \ldots, 1, 0, \ldots, 0]$ for the almost-disconnected components.

Spectral clustering

How best to divide a connected graph into two?

- Suppose the graph is almost disconnected.
- Then there is an eigenvector $[1, 1, \ldots, 1]$ and no other eigenvectors with eigenvalue 0.
- But there is another eigenvector that is very close to $[1, \ldots, 1, 0, \ldots, 0]$ for the almost-disconnected components.
- We can take the eigenvector v_2 with second-smallest eigenvalue.

Spectral clustering

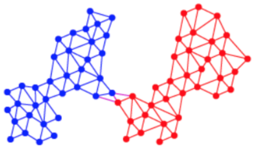
How best to divide a connected graph into two?

- Suppose the graph is almost disconnected.
- Then there is an eigenvector $[1, 1, \ldots, 1]$ and no other eigenvectors with eigenvalue 0.
- But there is another eigenvector that is very close to $[1, \ldots, 1, 0, \ldots, 0]$ for the almost-disconnected components.
- We can take the eigenvector v_2 with second-smallest eigenvalue.
- And compute $\mathrm{sign}(v_2) = (\pm 1, \pm 1, \ldots, \pm 1)$.

Spectral clustering

How best to divide a connected graph into two?

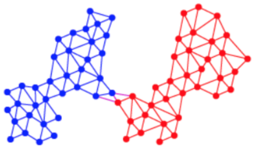
- Suppose the graph is almost disconnected.
- Then there is an eigenvector $[1, 1, \ldots, 1]$ and no other eigenvectors with eigenvalue 0.
- But there is another eigenvector that is very close to $[1, \ldots, 1, 0, \ldots, 0]$ for the almost-disconnected components.
- We can take the eigenvector v_2 with second-smallest eigenvalue.
- And compute $\text{sign}(v_2) = (\pm 1, \pm 1, \ldots, \pm 1)$.
- Those entries with +1 are in component 1, those with -1 are in component 2.



Spectral clustering

How best to divide a connected graph into two?

- Suppose the graph is almost disconnected.
- Then there is an eigenvector $[1, 1, \ldots, 1]$ and no other eigenvectors with eigenvalue 0.
- But there is another eigenvector that is very close to $[1, \ldots, 1, 0, \ldots, 0]$ for the almost-disconnected components.
- We can take the eigenvector v_2 with second-smallest eigenvalue.
- And compute $\text{sign}(v_2) = (\pm 1, \pm 1, \ldots, \pm 1)$.
- Those entries with +1 are in component 1, those with -1 are in component 2.



- Can formalize this by looking at min of $x^T L x$ for ± 1 vectors x.

Next time!

Calculus I