Again topic for a separate course...

SPECTRAL GRAPH THEORY

§ I

Study properties of graphs by exploring spectral properties (eigenvalues, eigenvectors) of matrices associated to graphs

Let us start with quick recap of facts and notation used in next few lectures

u, v vectors in $\mathbb{R}^n$ $\quad v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$

When it matters

Think of vectors as column vectors

$$\langle u, v \rangle = u^T \cdot v = \sum_{i=1}^{n} u_i v_i$$

u and v ORTHOGONAL, $u \perp v$, if $\langle u, v \rangle = 0$

ORTHONORMAL if also $\|u\|_2 = \|v\|_2 = 1$

L_2-norm of vector

$$\|v\|_2 = \sqrt{\langle v, v \rangle} = \sqrt{\sum_{i=1}^{n} v_i^2}$$

Will sometimes be sloppy and drop subscript 2 when clear from context

v unit vector if $\|v\|_2 = 1$

PYTHAGOREAN THEOREM

If $u \perp v$, then $\|u+v\|_2^2 = \|u\|_2^2 + \|v\|_2^2$

L_1 - norm $\quad \|v\|_1 = \sum_{i=1}^n |v_i|$
(sometimes denoted $|v|_1$)

L_p - norm $\quad \|v\|_p = \left(\sum_{i=1}^n |v_i|^p\right)^{1/p}$

All of these are indeed NORMS, i.e.:

1) $\|v\| \geq 0$ with equality iff $v = 0$
2) $\|\alpha v\| = |\alpha| \, \|v\| \qquad \alpha \in \mathbb{R}$
3) $\|u + v\| \leq \|u\| + \|v\|$

Useful to be able to switch between norms

PROPOSITION 1

For $v \in \mathbb{R}^n$ it holds that $\|v\|_1 / \sqrt{n} \leq \|v\|_2 \leq \|v\|_1$

Proof uses

CAUCHY-SCHWARZ INEQUALITY

$$|\langle u, v\rangle| \leq \|u\|_2 \, \|v\|_2$$

<u>Proof of Prop 1</u>

Let u be vector s.t. $u_i = |v_i|$

Let $\mathbb{1} = (1, 1, \ldots, 1)^T$

Then $|\langle \mathbb{1}, u\rangle| = \|v\|_1 \leq \|\mathbb{1}\|_2 \cdot \|u\|_2 = \sqrt{n}\,\|v\|_2$

so $\|v\|_1 / \sqrt{n} \leq \|v\|_2$.

$$\|v\|_1^2 = \left(\sum_{i=1}^n u_i\right)^2 = \sum_{i=1}^n u_i^2 + 2\sum_{i \neq j} u_i u_j \geq \sum u_i^2 = \|v\|_2^2$$

Taking square roots yields $\|v\|_2 \leq \|v\|_1$.

Consider undirected graphs $G = (V, E)$ S $\underline{III}$
Usually simple graphs (i.e., no loops or parallel edges) though sometimes helpful to allow loops and multiedges for technical reasons

Unless otherwise specified, will assume all graphs are d-regular (for some d) i.e., each vertex incident to d edges. Not superessential, but makes our life easier

Identify set of vertices with $[n]$ so that we can use vertices as integer indices

Some matrices associated to a graph

ADJACENCY MATRIX A_G

$$A_G(u,v) = \begin{cases} 1 & \text{if } (u,v) \in E \\ 0 & \text{otherwise} \end{cases}$$

NORMALIZED ADJACENCY MATRIX

$$\tilde{A}_G = \frac{1}{d} A_G$$

(sometimes drop tilde when normalization clear from context)

SIV

$\tilde{A}_G$ can be viewed as "diffusion matrix"
Each vertex starts out with some amount of "stuff" x_i; total distribution $x \in \mathbb{R}^n$
At each time step, stuff at vertex i distributed uniformly to neighbours. Then distribution after one time step is

$$\tilde{A}_G \, x$$

Spectral theory can provide understanding of this process.

LAPLACIAN MATRIX

$$\mathcal{L}_G = d \cdot I - A_G$$

$$\mathcal{L}_G(u,v) = \begin{cases} d & \text{if } u=v \\ -1 & \text{if } (u,v) \in E \quad (u \neq v) \\ 0 & \text{otherwise} \end{cases}$$

Defines nice quadratic form on graph

$$x^T \mathcal{L}_G \, x = \sum_{(u,v) \in E} (x_u - x_v)^2$$

[just expand LHS and think carefully]

Measures the smoothness of x viewed as a function $x : [n] \to \mathbb{R}$
[where $x(i) = x_i$; will be convenient to use x_i and $x(i)$ interchangeably at times]

NORMALIZED LAPLACIAN

$$\tilde{\mathcal{L}}_G = \frac{1}{d} \mathcal{L}_G$$

Choice of matrix to study not too important for regular graphs

Can make argument that Laplacian is "the right matrix" to study (properly generalized to also non-regular graphs)

But we won't need to go super-deep, so we will stick with <u>(normalized) adjacency matrix</u>

Recall: For $n \times n$ matrix A, (non-zero) v is an EIGENVECTOR with EIGENVALUE μ if

$$A v = \mu v$$

Note that all matrices we associated to graph G are <u>real</u> and <u>symmetric</u>.

Thanks to this, a miracle occurs

$$A = A^T$$

where A^T is the TRANSPOSE of A, i.e., matrix such that $A^T(i,j) = A(j,i)$

SPECTRAL THEOREM

Let A $n \times n$ symmetric real matrix.
Then A can be decomposed as

CAN BE VERY UNTRUE IF A NOT SYMMETRIC

$$A = \sum_{i=1}^{n} \mu_i \vec{u}_i \vec{u}_i^T = \begin{pmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} \mu_1 & 0 & & 0 \\ 0 & \mu_2 & & \\ & & \ddots & \\ 0 & & & \mu_n \end{pmatrix} \begin{pmatrix} - & u_1 & - \\ - & u_2 & - \\ & \vdots & \\ - & u_n & - \end{pmatrix}$$

where μ_i are the eigenvalues with associated eigenvectors $\vec{u}_i$ and $(\vec{u}_1, \ldots, \vec{u}_n)$ are <u>orthonormal</u>.

$\left(\text{so} \begin{pmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{pmatrix} \begin{pmatrix} - u_1 - \\ \vdots \\ - u_n - \end{pmatrix} = I \right)$

We don't need to prove this theorem, but can just happily use it.

However, just in case you are curious, let us remind ourselves how to prove parts of this.

FACT A A real symmetric matrix has real eigenvalues

<u>Proof</u> Let us look at complex numbers $z \in \mathbb{C}$

$$z = a + bi \quad , a, b \in \mathbb{R}, \; i^2 = -1$$

Conjugate $\quad z^* = a - bi$

$$z z^* = a^2 + b^2 > 0 \quad \text{if} \quad z \neq 0$$

$$(z_1 z_2)^* = z_1^* z_2^* \qquad \boxed{z \text{ real iff } z = z^*}$$

For $v = (v_1, \ldots, v_n)^T \in \mathbb{C}^n$

$v^* = (v_1^*, \ldots, v_n^*)^T$

$$\langle u, v \rangle = \sum u_i^* v_i$$

Straightforward to verify:

$$\langle Av, w \rangle = \langle v, A^T w \rangle \quad A \text{ real matrix}$$

$$\langle \lambda v, w \rangle = \lambda^* \langle v, w \rangle$$

$$\langle v, \lambda w \rangle = \lambda \langle v, w \rangle$$

Suppose λ eigenvalue with eigenvector v

$$\begin{aligned} \lambda^* \langle v, v \rangle &= \langle \lambda v, v \rangle \\ &= \langle Av, v \rangle \\ &= \langle v, A^T v \rangle \\ &= \langle v, Av \rangle \\ &= \langle v, \lambda v \rangle \\ &= \lambda \langle v, v \rangle \end{aligned}$$

But $\langle v, v \rangle \in \mathbb{R}$ and > 0 if $v \neq 0$

Hence $\lambda^* = \lambda$ and $\lambda \in \mathbb{R}$ as claimed

Eigenvalue λ iff $Av = \lambda v$ $\quad v \neq 0$

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

Characteristic polynomial $\det(\lambda I - A)$

Degree $n \implies n$ ~~roots~~ roots in $\mathbb{C}$

By Fact A, all roots in fact in $\mathbb{R}$

FACT B Eigenvectors of different eigenvalues
are orthogonal (so if all eigenvalues of A are distinct then A has an orthogonal eigenbasis)

Proof

Suppose $Av = \lambda v$, $Aw = \mu w$;
$v, w \in \mathbb{R}^n$ $\lambda, \mu \in \mathbb{R}$ $\lambda \neq \mu$

$$\begin{aligned}
\lambda \langle v, w \rangle &= \langle \lambda v, w \rangle \\
&= \langle Av, w \rangle \\
&= \langle v, A^T w \rangle \\
&= \langle v, Aw \rangle \\
&= \langle v, \mu w \rangle \\
&= \mu \langle v, w \rangle
\end{aligned}$$

which implies $\langle v, w \rangle = 0$ since $\lambda \neq \mu$.

Remains to prove: If λ eigenvalue of higher multiplicity $d > 1$ (root of char poly $\det(\lambda I - A)$ of multiplicity d), then dimension of

$$\{ v \in \mathbb{R}^n \mid Av = \lambda v \}$$

also has dimension d.

This we will not do.

Consider d-regular graph G and normalized adjacency matrix $\tilde{A}_G$

(can also look at A_G – same eigenvectors; eigenvalues factor d larger)

What can eigenvalues of $\tilde{A}_G$ tell us about G? A lot!

Order them $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_n$

Then:

1. Largest eigenvalue $\mu_1 = 1$ (eigenvector $\mathbb{1}$)

2. $\mu_2 < \mu_1$ iff G connected

3. Multiplicity of eigenvalue 1 = dimension of eigenspace spanned by corresponding vectors = # connected components of G

4. $\mu_n = -1$ iff G is bipartite

5. If G bipartite and μ is eigenvalue, then $-\mu$ also eigenvalue (of same multiplicity)

6. If $|\mu_n| \ll \mu_2$, then G does not have very large cuts (partition of vertices into two sets s.t. almost all edges go between two sets) In particular, "not close to bipartite"

7. If $\mu_1 \gg \max(|\mu_2|, |\mu_n|)$, then G has excellent connectivity properties (so called EXPANDER – will talk a lot about this)

Look at Laplacian again

$$x^T \mathcal{L}_G x = \sum_{(u,v)\in E} (x_u - x_v)^2$$

non-negative

So $\mathcal{L}_G$ positive semidefinite

Eigenvector $\mathbb{1}$ with eigenvalue 0

Laplacian more informative for irregular graphs (but we will only look at regular graphs)

Next eigenvalue (>0 for connected graph) known as FIEDLER VALUE

Eigenvector FIEDLER VECTOR

Useful heuristic for graph partitioning

Can also use Fiedler vector and next vector as coordinate system for drawing graphs (see very nice plots in lecture notes by Spielman)

But from now on focus on (normalized) adjacency matrix

see result along the lines of
Will ~~show~~ 6 next – just need one more tool.

Will then spend two lectures (probably) talking about 7

Properties 1 – 5 are left as (useful and highly recommended) exercises.

Eigenvalues of symmetric real matrices can be characterized in terms of Rayleigh quotients

<u>DEF 2</u> The RAYLEIGH QUOTIENT of $x \in \mathbb{R}^n$ with respect to the real, symmetric $n \times n$ matrix A is

$$\frac{x^T A x}{x^T x}$$

<u>PROPOSITION 3</u>

$$\mu_n \leq \frac{x^T A x}{x^T x} \leq \mu_1$$

<u>Proof</u> Let $u_1, \ldots, u_n \in \mathbb{R}^n$ be orthonormal basis of eigenvectors guaranteed by Spectral Theorem.

Write $\quad x = \sum \alpha_j u_j \qquad \alpha_i \in \mathbb{R}$

Then $$\frac{x^T A x}{x^T x} = \frac{\sum \mu_j \alpha_j^2}{\sum \alpha_j^2} \qquad (*)$$

since $x^T x = \langle x, x \rangle = \langle \sum \alpha_i u_i, \sum \alpha_j u_j \rangle = \sum \alpha_j^2$ by orthonormality.

Since $\mu_n \sum_1 x_j^2 \leq \sum \mu_j x_j^2 \leq \mu_1 \sum_1 x_j^2$ the proposition follows. □

COROLLARY 4

Let A real symmetric matrix and let x be a non-zero vector maximizing the Rayleigh quotient wrt A. Then x is an eigenvector with eigenvalue $\dfrac{x^T A x}{x^T x} = \mu_1$.

Proof Clearly, by the calculations in (*) in proof of Prop 3 it holds that if x eigenvector with eigenvalue λ, then $\dfrac{x^T A x}{x^T x} = \lambda$

So eigenvector for μ_1 yields Rayleigh quotient μ_1 which is max possible. If x maximizes Rayleigh quotient, then all coeff $x_j \neq 0$ have $\mu_j = \mu_1$, or else

$$\frac{\sum_1 \mu_j x_j^2}{\sum_1 x_j^2} < \mu_1 \cdot \frac{\sum_1 x_j^2}{\sum_1 x_j^2}$$

Hence x must be an eigenvector with value μ_1 □

Even if we don't know exact eigenvalues, Rayleigh quotients for well-chosen "test vectors" can be used to give bounds.

Recall that in a (regular) bipartite graph G we have $\mu_n = -1$, and G has a large independent set (~~at least~~ half of the vertices).

Let us show that if $\mu_n \gg -1$, then the largest independent set must be small.

THEOREM 5 (HOFFMAN)
Let G be a d-regular n-vertex graph and let μ_n be the smallest eigenvalue of $\tilde{A}_G$. Then the size $\alpha(G)$ of the largest independent set of G is

$$\alpha(G) \leq n \cdot \frac{-\mu_n}{1-\mu_n}$$

<u>Remark</u> For $\mu_n = -1$ we get $\alpha(G) \leq n/2$ which is tight. (Why?)
As μ_n gets closer to 0, the bound gets closer to $(-\mu_n) \cdot n$ (But will always be bounded away from 0)

<u>Proof</u> Let S be an independent set in G. Consider the vector x with value $n - |S|$ on coordinates $u \in S$ and value $-|S|$ on $v \notin S$. Then we can use the Rayleigh quotient to bound μ_n by

$$\mu_n \leq \frac{x^T \tilde{A}_G x}{x^T x}$$

We have

$$x^T x = |S|\,(n-|S|)^2 + (n-|S|)\,|S|^2$$

$$= |S|\,(n-|S|)\,[n-|S|+|S|]$$

$$= n\,|S|\,(n-|S|) \qquad (1)$$

for the denominator. The numerator is

$$x^T \tilde{A}_G\, x = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i\, a_{ij}\, x_j \qquad (\dagger)$$

for $\tilde{A}_G = (a_{ij})$. Only non-zero terms are for edges $(u,v) \in E$, where each edge gets counted twice as "directed edges" $(u \to v)$ and $(v \to u)$.

There are no edges in $S \times S$
There are $2d|S|$ directed edges between S and $\bar{S}$
Remaining $dn - 2d|S|$ edges are from S to S

So we get

$$x^T \tilde{A}_G\, x = \frac{1}{d}\left[|S|^2(nd - 2|S|d) - 2|S|d(n-|S|)|S|\right]$$

$$= \frac{1}{d}|S|^2\left(nd - 2|S|d - 2nd + 2|S|d\right)$$

$$= \frac{1}{d}\,|S|^2\,(-nd)$$

$$= -n|S|^2 \qquad (2)$$

Using (1) and (2) we get

$$\mu_n \leq \frac{x^T \tilde{A}_G x}{x^T x} = \frac{-n|S|^2}{n|S|(n-|S|)}$$

$$= \frac{-|S|}{n-|S|}$$

Rearranging we obtain

$$\mu_n \cdot n \leq (\mu_n - 1)|S|$$

$$\frac{\mu_n \cdot n}{\mu_n - 1} \geq |S|$$

How do we know $\mu_n - 1 \neq 0$, BTW?

or

$$|S| \leq n \cdot \frac{-\mu_n}{1-\mu_n}$$

and the theorem follows. $\boxed{\text{Ч}}$

RANDOM WALKS ON GRAPHS

Suppose we have graph $G = (V, E)$

Identify $V = [n]$

- Start at some vertex v
- At each time step, move to randomly chosen neighbour

What does probability distribution look like after t steps?

Does it get close to uniform distribution?

If so, how fast?

<u>Answer</u>: Eigenvalues of $\tilde{A}_G$ can tell us.

From now on, G understood from context and write simply A instead of $\tilde{A}_G$

Let $p \in \mathbb{R}^n$ initial probability distribution

$\forall i \; p_i \geq 0 \qquad \|p\|_1 = \sum_{i=1}^{n} |p_i| = \sum_i p_i = 1$

Let q distribution obtained by

- sampling start vertex with prob in p
- move to random neighbour

Probability to end up in j =

= pick neighbour i of j $\cdot$ move over edge (i,j)

$$= \cancel{\sum_{(i,j)\in E}} \sum_{\substack{i \in [n] \\ (i,j) \in E}} p_i \cdot \frac{1}{d} = \sum_{i=1}^{n} a_{ji} \, p_i$$

So distribution is $\quad q = A p$

After t steps distribution $A^t p$

Therefore A also called RANDOM WALK MATRIX

A is symmetric, all entries in $[0,1]$, and all rows and all columns sum to 1 (S XVI)

symmetric STOCHASTIC MATRIX

Recall eigenvalues denoted

$$\mu_1 \geq \mu_2 \geq \dots \geq \mu_n$$

Take absolute values and sort

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$$

$$\lambda_1 = 1$$

$$(*) \quad \lambda_2 = \max(|\mu_2|, |\mu_n|) \quad [= \max(\mu_2, -\mu_n)]$$

λ_2 also denoted $\lambda(G)$ (or just λ or sometimes $\lambda(A)$) sometimes (not quite correctly) called "second largest eigenvalue"

Using Rayleigh quotients, can convince ourselves that

$$\lambda(G) = \max_{x \perp \mathbb{1}} \left| \frac{x^T A x}{x^T x} \right|$$

(because of $(*)$ since eigenvectors corresponding to μ_2 and μ_n live in space $\{x \in \mathbb{R}^n \mid x \perp \mathbb{1}\} = \mathbb{1}^\perp$)

PROPOSITION 6 $\quad \lambda(G) = \max_{\substack{x \perp \mathbb{1} \\ \|x\|_2 = 1}} \|Ax\|_2$

Is another useful characterization

Proof Write in eigen basis $x = \sum \alpha_i u_i$

Recall all u_i orthogonal note $\alpha_1 = 0$

$$\|Ax\|_2^2 = \left\| \sum_{i=2}^{n} \alpha_i \mu_i u_i \right\|_2^2$$

$$= \sum_{i=2}^{n} |\mu_i|^2 \|\alpha_i u_i\|_2^2$$

$$\leq \lambda(G)^2 \sum_{i=2}^{n} \|\alpha_i u_i\|_2^2$$

$$= \lambda(G)^2 \|x\|_2^2$$

Plugging in eigenvector u_2 or u_n will clearly yield equality.

DEF 7 The SPECTRAL GAP of G is

$$1 - \lambda(G)$$

If G has a large spectral gap, then random walk on G quickly approaches uniform distribution

Described by $\mathbb{1}/n = (1/n, 1/n, \dots, 1/n)^T$

THM 8 Let G be any n-vertex d-regular graph and let p be any initial probability distribution. Then the distribution of the random walk after l steps $A^l p$ satisfies

$$\| A^l p - \mathbb{1}/n \| \leq \lambda(G)^l$$

<u>Proof</u> We just saw

$$\|Av\|_2 \leq \lambda \|v\|_2$$ (write $\lambda = \lambda(A)$ below)

for all $v \perp \mathbb{1}$.

If $v \perp \mathbb{1}$, then $Av \perp \mathbb{1}$ so A maps $\mathbb{1}^\perp$ to itself. To see this:

$$\langle \mathbb{1}, Av \rangle = \langle A^T \mathbb{1}, v \rangle$$
$$= \langle A\mathbb{1}, v \rangle$$
$$= \langle \mathbb{1}, v \rangle = 0.$$

Hence $\|A^2 v\|_2 \leq \lambda \|Av\|_2 \leq \lambda^2 \|v\|_2$
and

$$\lambda(A^\ell) \leq \lambda^\ell(A)$$

[In fact, can show that eigenvalues of A^ℓ are μ_i^ℓ, $i = 1, \dots, n$, so $\lambda(A^\ell) = (\lambda(A))^\ell$]

Write $p = \alpha(\mathbb{1}/n) + p'$

for $p' \perp \mathbb{1}$ $\quad \sum p_i' = 0$ }
p probability distr } $\alpha = 1$

$$A^\ell p = A^\ell(\mathbb{1}/n + p') = \mathbb{1}/n + A^\ell p' \quad (\dagger)$$

Pythagoras:

$$\|p\|_2^2 = \|\mathbb{1}/n + p'\|_2^2 = \|\mathbb{1}/n\|_2^2 + \|p'\|_2^2$$

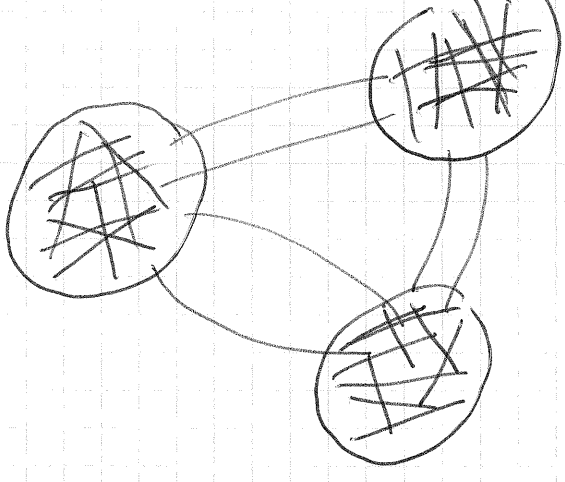
so $\|p'\|_2 \leq \|p\|_2$
$\leq \|p\|_1$ (L_1-norm larger)
$= 1$ (p prob distr)

Using (†) we obtain

$$\| A^{\ell} p - \mathbb{1}/n \|_2 = \| A^{\ell} p' \|_2$$
$$\leq \lambda^{\ell} \|p'\|_2$$
$$\leq \lambda^{\ell}$$

Graphs with large spectral gap have to be <u>well-connected</u>

Cannot look like



dense components with few outgoing edges, because then random walk would not mix quickly.

We will talk about such well-connected graphs next (and for a while):

EXPANDER GRAPHS

Next couple of lectures based on material from Chapter 21 in Arora-Barak – might be helpful as side reading