Lecture 14 - Spectral Graph Theory II

Recap:

- G: undirected graph (self-loops, parallel edges OK; think of G as regular, for a warmup)
  - $\{f: V \to \mathbb{R}\}$ is an n-dim. vector space $(n = |V|)$
- "weighted" inner product $\langle f, g \rangle_\pi = \mathop{E}_{u \sim \pi}[f(u)g(u)]$
- π is invariant/stationary distribution on V:
  - choose unif. random edge $u \sim v$ $(\equiv v \sim u)$
  - output u
- $\pi(u) = \frac{\deg(u)}{2m}$ is proportional to $\deg(u)$ $(m = |E|)$

  $\to$ uniform distribution on V if G regular

- If $u_0 \sim \pi$ and then $u_0 \to u_1 \to u_2 \to \dots \to u_t$ is a standard rand. walk, then each u_i is $\sim \pi$.
- If $f: V \to \{0,1\}$ is $f = 1_S$, $S \subseteq V$, ("fractional volume, $|S|/|V|$ if G is regular")

  "$\|f\|_2^2$" $= \langle f, f \rangle = \mathop{E}_{u \sim \pi}[f(u)^2] = \Pr_{u \sim \pi}[u \in S] =$ "$\mathrm{vol}(S)$"

- Quadratic form $\mathcal{E}[f] := \frac{1}{2} \mathop{\mathbb{E}}_{u \sim v} \left[(f(u) - f(v))^2 \right]$

  Compare: $\mathrm{Var}_{\pi}[f] = \frac{1}{2} \mathop{\mathbb{E}}_{\substack{u \sim \pi \\ v \sim \pi \\ \text{indep.}}} \left[(f(u) - f(v))^2 \right]$

- $\mathcal{E}[\mathbb{1}_S] = \Pr_{u \sim v} [u \in S, v \notin S]$ (prob. a random edge steps out of $S \subseteq V$)

$\min_f \mathcal{E}[f]$? If G has connected components $S_1, \ldots, S_\ell$ then $\mathbb{1}_{S_1}, \ldots, \mathbb{1}_{S_\ell}$ are linearly independent and their span is all f s.t. $\mathcal{E}[f] = 0$.

$\max_f \mathcal{E}[f]$ s.t. $\|f\|_2^2 = \langle f, f \rangle = \mathop{\mathbb{E}}_{u \sim \pi} [f(u)^2] \leq 1$?

(We need some scaling condition, as $\mathcal{E}[c \cdot f] = c^2 \cdot \mathcal{E}[f]$.)

(equiv.: s.t. $\mathbb{E}[f^2] = 1$ or, s.t. $\mathrm{Var}[f] \leq 1$ ($=$))

fact: $\mathcal{E}[f] \leq 2 \|f\|_2^2$, with equality possible iff G is bipartite.

// end recap

We'll look some more at max/min of $\mathcal{E}[f]$ with an eye toward how it relates to "mixing time" for rand. walk on G.

Let's recap last proof...

Proof: $\mathcal{E}[f] = \text{arithmetic...} = \|f\|_2^2 - \boxed{\mathop{E}_{u\sim v}[f(u)f(v)]}$

$\leq \|f\|_2^2$ by Cauchy-Schwarz.

The boxed quantity is interesting to study. Equals...

$$\square = \mathop{E}_{u\sim\pi}\mathop{E}_{v\sim u}[f(u)\cdot f(v)] = \mathop{E}_{u\sim\pi}\left[f(u)\cdot \mathop{E}_{v\sim u}[f(v)]\right]$$

$v\sim u$: a rand. neighbor of u

(this is a function of u) → average val. of f on u's neighbors

def: $Kf : V \to \mathbb{R}$, $(Kf)(u) = \mathop{E}_{v\sim u}[f(v)]$

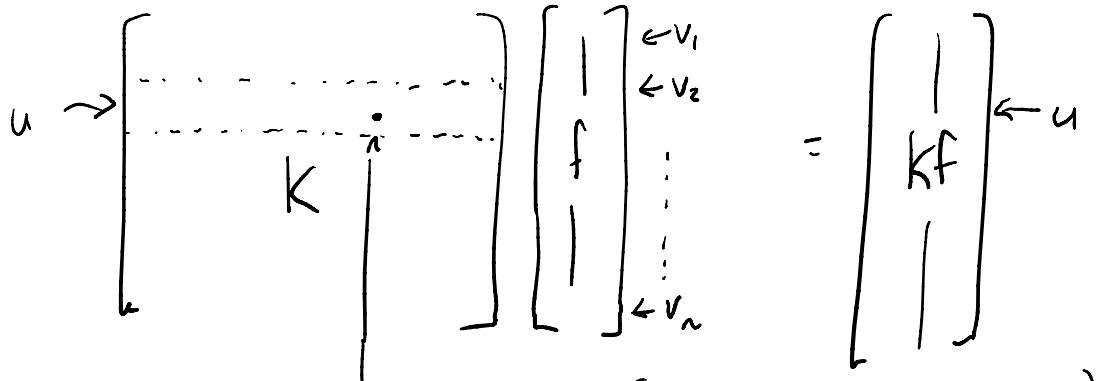
We write it like "Kf" because we can think of "K" as a (linear) operator on functions...

def: K is Markov / transition / normalized adjacency operator (matrix) for G.

obs: K is a "linear operator": $K(f+g) = Kf + Kg$

proof: Linearity of expectation

(It's a linear operator on fcns $f \equiv$ vectors f.
Hence it corresponds to a matrix.)



$$K[u,v] = \begin{cases} \frac{1}{\deg(u)} & \text{if } (u,v) \in E \\ 0 & \text{else} \end{cases}$$

$$= \Pr[u \to v \mid @u]$$

(in one step of S.R.W.)

$\therefore$ K is adjacency matrix A normalized to have row-sum 1 (it's a "stochastic" matrix)

- If G is d-regular, $K = \frac{1}{d} \cdot A$ and

K is symmetric $(K^T = K)$

fact: For $f, g : V \to \mathbb{R}$,

$$\langle f, Kg \rangle = \mathop{E}_{u \sim v} [f(u) \cdot g(v)]$$

(Should almost think of this like the def. of K.)

cor: $\because u \sim v \equiv v \sim u$, also equals

$$\mathop{E}_{u \sim v} [g(u) \cdot f(v)] = \langle Kf, g \rangle.$$

"K is self-adjoint"

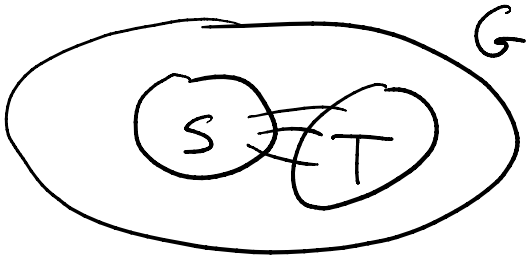
(The appropriate generalization of "K is symmetric" to the irregular G case.)

Proof:

$$\begin{aligned} \langle f, Kg \rangle &= \mathop{E}_{u \sim \pi} [f(u) \cdot (Kg)(u)] \\ &= \mathop{E}_{u \sim \pi} \left[f(u) \cdot \mathop{E}_{v \sim u} [g(v)] \right] \\ &= \mathop{E}_{\underbrace{u \sim \pi, v \sim u}_{u \sim v}} [f(u) \cdot g(v)] \end{aligned}$$

■

e.g.:



$S, T \subseteq V$ (don't have to be disjoint)

Let $f = \mathbb{1}_S$, $g = \mathbb{1}_T$.

$$\langle f, Kg \rangle = \mathop{E}_{u \sim v} \left[\mathbb{1}_S(u) \cdot \mathbb{1}_T(v) \right]$$

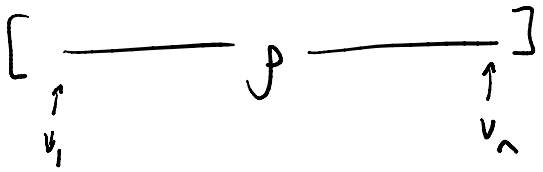
$$= \mathop{Pr}_{\substack{u \sim v \\ \text{unif. rand edge}}} \left[u \in S, v \in T \right].$$

(It's prob. of a rand. edge taking you from S to T in one step.)

(Let's now begin to think about taking a multiple-step random walk, and how this relates to K.)

obs.: Let ρ be any probability distrib. on V.

Think of it as a row vector:



Say we do:
- Draw $u \sim \rho$
- Do one random step $u \to v$.

Let ρ' be probability distrib. on V.

Then (exercise) $[\rho]\begin{bmatrix} & K & \end{bmatrix} = [\rho']$

cor: $\pi K = \pi$ (as π is the invariant/stationary distribution)

(Now if we start in distribution ρ and do multiple steps, corresponds to multiple multiplications by K.)

Claim: $K^2 = K \circ K$ operates as

$$(K^2 f)(u) = \mathop{\mathbb{E}}_{\substack{u \to w \\ \text{2 steps}}} [f(w)]$$

Proof: Let $g = Kf$. $K^2 f = K(Kf) = Kg$, so

$$(K^2 f)(u) = (Kg)(u) = \mathop{\mathbb{E}}_{v \sim u} [g(v)] = \mathop{\mathbb{E}}_{v \sim u} [(Kf)(v)]$$

$$= \mathop{\mathbb{E}}_{v \sim u} \left[\mathop{\mathbb{E}}_{w \sim v} [f(w)] \right]. \blacksquare$$

cor: $\forall t \in \mathbb{N}$, $(K^t f)(u) = \mathop{\mathbb{E}}_{\substack{u \to w \\ t \text{ steps} \\ \text{of S.R.W.}}} [f(w)]$

(K^t: matrix power, or t-fold composition)

Even $t = 0$ ✓ $(I \cdot f)(u) = \mathop{\mathbb{E}}_{w := u} [f(w)] = f(u)$. ✓

Back to $\mathcal{E}[f] = \frac{1}{2} \mathop{E}_{u \sim v} [(f(u) - f(v))^2]$

(by simple calc. →) $= \langle f, f \rangle - \mathop{E}_{u \sim v} [f(u) f(v)] \rightarrow \langle f, Kf \rangle$

$= \langle f, If - Kf \rangle$

$= \langle f, \underbrace{(I - K)}_{L} f \rangle$

(the all-important "Laplacian")

def: $L = I - K$ is (normalized) Laplacian operator.

(d-regular case: $L = I - \frac{1}{d} A = \frac{1}{d}(\underbrace{dI - A}_{\text{"unnormalized Laplacian"}})$)

(in irregular case, those who don't like $\langle , \rangle_\pi$ have a tortured definition involving "$D^{-1/2}$")

I.e., $Lf : V \to \mathbb{R}$ is $Lf(u) = f(u) - \mathop{E}_{v \sim u} [f(v)]$.

Q: What is "meaning" of L?

A: It's the operator s.t. $\langle f, Lf \rangle = \mathcal{E}[f]$

the awesomely meaningful Quadratic Form

Let's see how this definition relates to an important algorithms problem, "Sparsest Cut".

Let $S \subseteq V$, let $f = \mathbb{1}_S$.

$$\langle f, Lf \rangle = \mathcal{E}[f] = \Pr_{u \sim v}[u \in S, v \notin S]$$

(= "volume" of $\vec{\partial} S$: edges leaving S)

$$\langle f, f \rangle = \mathop{E}_{u \sim \pi}[f(u)^2] = \Pr_{u \sim \pi}[u \in S]$$ (= "volume" of S)

$$\therefore \text{ratio } \frac{\langle f, Lf \rangle}{\langle f, f \rangle} = \Pr_{u \sim v}[v \notin S \mid u \in S]$$

↑ conditional probability

$$= \Pr[\text{pick rand } u \sim S \text{ (proport. to deg.), do 1 step, that you go outside } S]$$

$\in [0,1]$: "escape prob. of S"

def: $\Phi[S]$ = conductance of S = ↖ this quantity

def: Sparsest Cut: Given G, find S of $\text{vol}(S) \leq \frac{1}{2}$ minimizing $\Phi[S]$. (A very important algorithmic task, esp. for divide & conquer on graphs. Approximability is wide open.)

Connections to spectral graph theory are a little easier when considering maximizing $\mathcal{E}[f]$, cf. Max-Cut. So let's see connections to eigenvalues...

$$\max \mathcal{E}[f] = \langle f, Lf \rangle = \tfrac{1}{2} \mathop{\mathbb{E}}_{u \sim v} \left[(f(u) - f(v))^2 \right]$$

↖ continuous fcn. of f's vals

$$\text{s.t. } \|f\|_2^2 = \langle f, f \rangle = \mathop{\mathbb{E}}_{u \sim \pi} [f(u)^2] = 1$$

↖ compact set in $\mathbb{R}^n$ (an ellipsoid)

$\therefore$ a maximizer exists, call it φ.

Claim: $L\varphi = \lambda \cdot \varphi$ for some $\lambda \in \mathbb{R}$ (i.e., $L\varphi$ is parallel to φ)

(We'll prove the claim shortly; let's assume it for now.)

cor: $\mathcal{E}[\varphi] = \langle \varphi, \lambda \varphi \rangle = \lambda \cdot \|\varphi\|_2^2 = \lambda$.

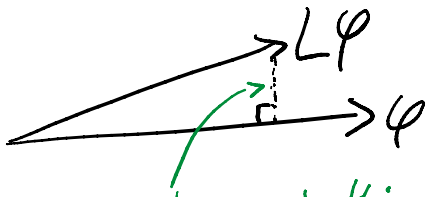
← we know $0 \le \lambda \le 2$ (↑ iff G bipartite)

fact: $\mathop{\mathbb{E}}_{u \sim \pi} [\varphi(u)] = 0 \iff \langle \varphi, \mathbb{1} \rangle = 0$.

Pf: Else replace φ with $\varphi - \mathbb{E}[\varphi]$. Doesn't change $\mathcal{E}[\varphi]$, but reduces $\|\varphi\|_2^2$. (Equiv., as we saw last time, maximizer will have $\mathbb{E}[\varphi^2] = \mathrm{Var}[\varphi] = 1$.)

Proof of claim: ("Lagrange Multipliers")

AFSOC $L\varphi$, φ not parallel:



Let ψ be unit vec. in this direction.

$\langle \psi, \varphi \rangle_\pi = 0$

Let $\varepsilon \neq 0$ be small. Let $f = \varphi + \varepsilon \cdot \psi$.

$$\|f\|_2^2 = 1 + \varepsilon^2 \quad \text{(Pythagorus)}$$

$$\langle f, Lf \rangle = \langle \varphi + \varepsilon\psi,\ L\varphi + \varepsilon L\psi \rangle$$

$$= \langle \varphi, L\varphi \rangle + 2\varepsilon \langle \psi, L\varphi \rangle + O(\varepsilon^2)$$

(used L is self-adjoint)

$$\therefore \frac{\langle f, Lf \rangle}{\langle f, f \rangle} = \langle \varphi, L\varphi \rangle + 2\varepsilon \underline{\langle \psi, L\varphi \rangle} + O(\varepsilon^2)$$

$\neq 0$, by picture

$\langle f, f \rangle = 1 + \varepsilon^2$

$\therefore$ can make this $> \langle \varphi, L\varphi \rangle$ by varying ε, contradicting optimality of φ.

∎

Great. So now that we have maximizing direction φ, let's disallow it — restrict to its orthogonal subspace — and ask for new maximizer.

Consider: $\max \langle f, Lf \rangle$
s.t. $\langle f, f \rangle = 1$, $f \perp \varphi$
(still a compact set)

Same argument $\Rightarrow$ maximizing φ' has $L\varphi' = \lambda' \varphi'$, $\lambda' \leq \lambda$,
and $E[\varphi'] = 0 \iff \varphi' \perp \mathbb{1}$ (and $\perp \varphi$),

Now max over $f \perp \varphi, \varphi'$, etc. ...
(All maximizers will be $\perp \mathbb{1}$, so last one will be $\mathbb{1}$.)

Theorem ("Spectral Theorem"): Given G, $\exists$ orthonormal functions ("eigenfunctions"/"eigenvectors") $\varphi_0, \varphi_1, \ldots, \varphi_n$
(orthonormal means $\langle \varphi_i, \varphi_j \rangle = \begin{cases} 1 \text{ if } i=j \\ 0 \text{ else} \end{cases}$)
with $\varphi_0 \equiv 1$ (these $\varphi_i : V \to \mathbb{R}$ form a basis)
& real numbers ("eigenvalues") $0 \leq \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{n-1} \leq 2$
such that $L\varphi_i = \lambda_i \varphi_i \quad \forall 0 \leq i \leq n-1$. ($\varphi_{n-1}, \lambda_{n-1}$ is argument's first (φ, λ) pair.)