Math 312, lecture 18

Last time: Norms on vector spaces, maps

Fix $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, $V = \mathbb{F}$-vsp

A norm on V is a map $\|\cdot\| : V \to \mathbb{R}_{\geq 0}$ s.t.

(1) $\|\underline{u} + \underline{v}\| \leq \|\underline{u}\| + \|\underline{v}\|$

(2) $\|\alpha \underline{v}\| = |\alpha|_{\mathbb{F}} \cdot \|\underline{v}\|$ $\qquad (\|\underline{0}\| = 0)$

(3) $\|\underline{v}\| = 0 \Rightarrow \underline{v} = \underline{0}$.

Examples: $\|\underline{x}\|_2 = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2}$ $\quad \underline{x} \in \mathbb{F}^n$

f cts on K cpt, $\|f\|_\infty = \max\{|f(x)| : x \in K\}$

(vsp is $C(K)$)

If U, V norms vsp, $T : U \to V$ linear,

T cts iff T is bounded: $\|T\underline{u}\|_V \leq M \cdot \|\underline{u}\|_U$ for some $M \geq 0$ indep of $\underline{u}$

Set $\|T\|_{U \to V} = \inf\left\{ \begin{matrix} \text{bounds} \\ M \end{matrix} \right\} = \sup\left\{ \frac{\|T\underline{u}\|_V}{\|\underline{u}\|_U} : \underline{u} \neq \underline{0} \right\}$

Lemma: $\|\cdot\|_{V\to V}$ is a norm on subspace $\mathrm{Hom}_b(V,V)$ of bounded linear maps.

Today: Application: Power Method

Problem: Given $T \in \mathrm{End}_F(V)$, want to find the eigenvalues (and eigenvectors).

- If A is symmetric/Hermitian can find all eigenvalues, eigenvectors in time $O(n^3)$

- Today: only want one (or a few) eigenvalues/eigenvectors

Application: PageRank.

Problem: want to find web page "most relevant" to a search query.

Model: WWW is a directed graph: has nodes V (webpages), edges $\mathcal{E} = \{(x_i, y_i)\} \subset V \times V$ hyper links

random walk on web is a random process where if at time t we are at page X_t, we click on an outgoing link, uniformly at random

$$\Pr\{X_{t+1} = y \mid X_t = x\} = \begin{cases} 0 & (x,y) \notin \mathcal{E} \\ \frac{1}{d_x} & (x,y) \in \mathcal{E} \end{cases}$$

$$d_x = \#\{y \mid (x,y) \in \mathcal{E}\}$$

---

let $W = \mathbb{R}^V = \{f : V \to \mathbb{R}\}$

if $f \in W$

$$(Af)(x) = \frac{1}{d_x} \sum_{y \leftarrow x} f(y)$$

$$A\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix} = \begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix} \Rightarrow 1 \text{ is an eigenvalue!}$$

$$|Af(x)| = \frac{1}{d_x} \Big| \sum_{(x,y)\in\mathcal{E}} f(y) \Big| \leq \frac{1}{d_x} \sum_{(x,y)\in\mathcal{E}} |f(y)| \leq \|f\|_\infty$$

$$\Rightarrow \|Af\|_\infty \leq \|f\|_\infty \Rightarrow \|A\|_{\ell^\infty \to \ell^\infty} \leq 1.$$

(actually $\|A\|_\infty = 1$) $\Rightarrow$ all ev. satisfy $|\lambda| \leq 1$

(if $Af = \lambda f$, take norms on both sides)

Assume graph (V, E) connected, then $\lambda = 1$
$\Rightarrow f$ constant.

Expect 1 simple eigenvalue, all other e.v. satisfy $|\lambda| < 1$

Let A act on row vectors:

$\pi \in W^*$, what is πA ?

$$(\pi A)_y = \sum_x \pi(x) A_{xy}$$

interpretation: if π is a rule for randomly choosing a vertex x, πA is also a rule for randomly choosing a vertex by (1) choose from π
(2) do one step of RW along A.

check $\langle \pi A, \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \rangle = \langle \pi, A \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \rangle = \langle \pi, \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \rangle$

so total prob $\sum_x \pi(x)$ is conserved.

Know: have eigenvector of e.v. 1, i.e. $\tilde{\pi}$ s.t.

$$\tilde{\pi} A = \tilde{\pi}$$ stationary distribution

$\Rightarrow$ If we do more & more steps of RW (from any starting point), the prob of being at any place converges to $\tilde{\pi}$:

$$\pi_0 A^k \xrightarrow[k \to \infty]{} \tilde{\pi}$$

Better model: proportion $1-\varepsilon$ of time click on random link, proportion ε of time we return to starting page

Now RW operator is $Bf = (1-\varepsilon) A + \varepsilon \delta_{x_0} \cdot f$

$$(Bf)(x) = (1-\varepsilon) \frac{1}{d_x} \sum_{x \sim y} f(y) + \varepsilon \cdot f(x_0)$$

(still $B \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$, still has stationary distr. $\tilde{\pi}$)

<u>Def</u>: (Brin-Page) The Page Rank of x is $\tilde{\pi}(x)$.

(For more reading: spectral graph theory, Perron-Frobenius eigenvalue/eigenvector)

Back to LA: Have matrix B, largest e.v. λ

want to find λ, and an eigenvector.

Idea: ("Power method") if v any vector, in $A\underline{v}$, the component along largest eigenvector is enhanced. In $A^k \underline{v}$, enhances a lot.

Basic power method

let $\underline{v} = \underline{v}_0$ be any vector, consider $A^k \underline{v}$

Fix eigen basis $\{\underline{u}_i\}_{i=1}^n$: $A\underline{u}_i = \lambda_i \underline{u}_i$

$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots$

spectral gap

Then if $\underline{v} = \sum_{i=1}^n a_i \underline{u}_i$ has

$$A^k \underline{v} = \sum_{i=1}^{n} a_i \lambda_i^k \underline{u}_i$$

$$= a_1 \lambda_1^k \left(\underline{u}_1 + \sum_{i=2}^{n} \left(\frac{a_i}{a_1}\right) \cdot \left(\frac{\lambda_i}{\lambda_1}\right)^k \underline{u}_i \right)$$

By assumption $\left|\frac{\lambda_i}{\lambda_1}\right| \leq \left|\frac{\lambda_2}{\lambda_1}\right| \leq \rho < 1$

So $\left\| \sum_{i=2}^{n} \left(\frac{a_i}{a_2}\right)\left(\frac{\lambda_i}{\lambda_1}\right)^k u_i \right\| \leq \rho^k \cdot \left(\frac{\max|a_i|}{a_1}\right) \cdot \max_i \|u_i\|$

as $k \to \infty$ this decay exponentially

$$\Rightarrow \left| \left\| \underline{u}_1 + \sum_{i=2}^{n} * \right\| - \|\underline{u}_1\| \right| \xrightarrow[k \to 0]{} 0 \text{ exp}$$

$$\Rightarrow \|A^k \underline{v}\| \sim |a_1| \, |\lambda_1|^k \, \|\underline{v}_1\|$$

$$\Rightarrow \frac{A^k v}{\|A^k v\|} = (1 + o(1)) \cdot (\underline{v}_1 + \text{small})$$

- practically don't want to mult by A^k then normalize, want to normalize at each step
- don't want to compute A^k

Algorithm: Given square matrix A

(1) choose vector v_0

(2) repeatedly set $v_{j+1} = \frac{Av_j}{\|Av_j\|}$

(3) When $Av_k \approx \lambda v_k$ stop

Ex: $v_j = \frac{A^j v_0}{\|A^j v_0\|}$ (induction)

- Only do matrix-vector mult Av_j; if A is sparse, this is efficient
- Numerical errors aren't a huge deal: every mult by A enhances u_1.
- Rate of convergence depends $1-p = \frac{\lambda_1 - \lambda_2}{\lambda_1}$

really on ~~sep~~ $\lambda_1 - \lambda_2$.

- Algorithm projecting v on λ_1-eigenspace
- If λ_2 close, numerically eigenspace 2d

Ex: Extend to non-diagonable matrices
(still assume λ_1 simple, $|\lambda_2|<|\lambda_1|$)
Hint: use Jordan form, check if N Jordan block N^k grows at most polynomially in k, suppression by ρ^k still works

More eigenvalues: subspace methods

Suppose A is symmetric / Hermitian, use ℓ^2 norm

Instead of looking for $(\underline{u}_1, \lambda_1)$, want $\{(\underline{v}_i, \lambda_i)\}_{i=1}^{r}$
r fixed $(n \to \infty)$

Natural idea: V matrix with r orthogonal columns of length n
iteratively mult AV, orthogonalize columns (use "QR factorization").

· Better method: Lanczos method

Specific eigenvalues: inverse iteration

A still Hermitian, want e.v. closest to σ.
$\Leftrightarrow$ smallest e.v. of $A-\sigma$
$\Leftrightarrow$ largest e.v. of $(A-\sigma)^{-1}$!

"Algorithm: $v_{j+1} = \dfrac{(A-\sigma)^{-1} v_j}{\|(A-\sigma)^{-1} v_j\|}$."

really: set w_{j+1} by $(A-\sigma) w_{j+1} = v_j$

set $v_{j+1} = \dfrac{w_{j+1}}{\|w_{j+1}\|}$.

solve system of linear equations

- at initially don't need exact solutions to LA
- can apply this to subspace iteration, find r eigenvalues closest to σ.