March 18, 2021 CPSC 531F

- Problems up to now + a few added today & tomorrow = Homework 2

- From here $\left\{\begin{array}{l}\text{Homework 3}\\ \text{Presentation \& Notes}\end{array}\right.$

for Presentations – $\left\{\begin{array}{l}\text{email me}\\ \text{speak to me}\end{array}\right.$

- Homework – cite other references, but write out any proofs in your own words (in class notation) rather than just cite a theorem

- Homework: "Write down", "give" (also explain or justify unless explicitly saying not to)
- If you want, hand in HW 2 for feedback, up to 2 weeks from today.
- Hand in all homework 2 weeks after last problem assigned.
- Most recent problems: I gave ≈ 8 problems, asked for at least 5

Last time:

"Variation Principle"

We had inner product space $\bar{V}$,

($\mathbb{R}$ = scalars, also $\mathbb{C}$ = as scalars

would also work).

No harm in thinking of $\bar{V} = \mathbb{R}^n$,

but everything goes through for

n-dimensional vector spaces,

with inner product: $(\, , \,)$

or $(\, , \,)_{\bar{V}}$

Formally $(\, , \,)_{\bar{V}}$ is a map

$\bar{V} \times \bar{V} \to \mathbb{R}$, write $(\vec{v}_1, \vec{v}_2)_{\bar{V}}$

$\vec{v}_1 \quad \vec{v}_2 \mapsto (\vec{v}_1, \vec{v}_2)_{\bar{V}}$.

E.g.

(1) $(\vec{v}_1, \vec{v}_2) = \vec{v}_1 \cdot \vec{v}_2$

(2) Markov matrices, P, reversible!

$$(\vec{u}, \vec{v})_{\vec{\pi}} = \sum_{i=1}^{n} u_i v_i \pi_i$$

and

$$(P\vec{u}, \vec{v})_{\vec{\pi}} = (\vec{u}, P\vec{v})_{\vec{\pi}}$$

we say P is self-adjoint

w.r.t. inner product $(\, , \,)_{\vec{\pi}}$,

where $\vec{\pi}$ = stationary distribution

$$R_{\mathcal{L}}(\vec{v}) = \frac{(\mathcal{L}\vec{v}, \vec{v})}{(\vec{v}, \vec{v})}$$

$\mathcal{L} : \bar{V} \to \bar{V}$ (e.g. $\bar{V} = \mathbb{R}^n$,

$\mathcal{L}(\vec{v}) = A\vec{v}$ for some $A \in M_n(\mathbb{R})$)

Step 1: Imagine that $R_{\mathcal{L}}(\vec{v})$

has its maximum at $\vec{v}^*$

<u>Step 2</u>: We take $\vec{w} \in \mathbb{R}^n$

$$\vec{v}_\varepsilon = \vec{v}^* + \varepsilon \vec{w}$$

($\varepsilon \in \mathbb{R}$, think of ε small)

$$f(\varepsilon) = R_{\mathcal{L}}(\vec{v}_\varepsilon) :$$

~~by~~ by assumption $\varepsilon = 0$ is a maximum of $f(\varepsilon)$, i.e.

$f'(0) = 0$ (assuming f is differentiable near $\varepsilon = 0$)

$$R_{\mathcal{L}}(\vec{v}_\varepsilon) = C_0 + C_1\varepsilon + C_2\varepsilon^2 + \ldots$$

$$= \underset{\substack{\uparrow \\ R_{\mathcal{L}}(\vec{v}^*)}}{C_0} + \underset{\substack{\uparrow \\ C_1 = \text{messy}\ldots}}{C_1\varepsilon} + \text{order}(\varepsilon^2)$$

but $C_1 = 0$!!

Rayleigh quotient - diagonalizing the Δ = Laplace operator

Then!

(1) The variational consideration shows

$$\underbrace{\mathcal{L}(\vec{v}^{*})}_{\text{this had to orthogonal to any vector orthogonal to } \vec{v}^{*}} = c_1 \vec{v}^{*}$$

Also $\vec{v}^{*} \longrightarrow \vec{v}_1$

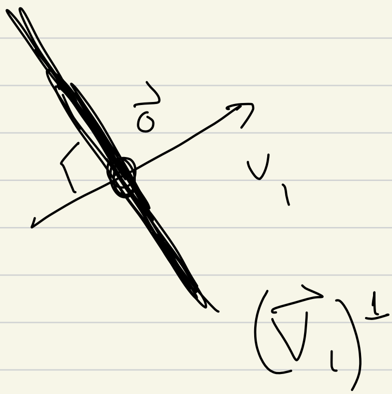
$$c_1 \longrightarrow \lambda_1 = R_{\mathcal{L}}(\vec{v}_1)$$

Now we get $\vec{v}_2$ ...

to get intuition, think of $\vec{v} \cdot \vec{v}_1$

Consider

$$(\vec{v}_1)^{\perp} = \{ \vec{v} \in \bar{V} \mid \overbrace{(\vec{v}, \vec{v}_1)}=0 \}$$



for $\cdot$ product

Imagine $\vec{v}_2$ maximizes

$R_{\mathcal{L}}(\vec{v})$ over all $\vec{v} \in (\vec{v}_1)^{\perp}$

$$(\vec{v}_1)^{\perp} \cap \{ \vec{v} \mid \underbrace{v_1^2 + \dots + v_n^2 = 1}^{\text{unit vectors}} \}$$

this set is also (sequentially) compact (i.e., closed and bounded),

Similar calculation!

if $\vec{w} \perp \vec{v}_1$ and $\vec{v}_2$

then consider!

$\mathcal{R}_{\mathcal{L}}(\vec{v}_2 + \varepsilon \vec{w})$ has max at $\varepsilon = 0$

$$\Rightarrow \quad \underbrace{\mathcal{L}(\vec{v}_2)}_{\text{this is orthogonal to any such } \vec{w}} = \underbrace{\tilde{c}_1 \vec{v}_1} + \tilde{c}_2 \vec{v}_2$$

then $\tilde{c}_1 = 0$

$$(\mathcal{L}(\vec{v}_2), \vec{v}_1) = (\vec{v}_2, \mathcal{L}\vec{v}_1)$$

$$= (\vec{v}_2, \lambda \vec{v}_1)$$

$\mathcal{L}$ is self-adjoint

$$= \lambda(\vec{v}_2, \vec{v}_1) = 0$$

$$\mathcal{L}(\vec{v}_2) = \hat{\tilde{c}}_2 \vec{v}_2$$

$$= \lambda_2 \vec{v}_2 \quad \text{and} \; ---$$

Now look at

$$\underbrace{(\vec{v}_1, \vec{v}_2)^\perp = \{\vec{v} \in \vec{V} \mid \vec{v} \text{ is orthog to both } \vec{v}_1, \vec{v}_2\}}$$

consider max $R_{\mathcal{L}}(\vec{v})$ over $\}$,

imagine its maximum is

attained at $\vec{v}_3$, etc...

This gives us

$$\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$$

with $\mathcal{L}(\vec{v}_i) = \lambda_i \vec{v}_i$

and

$\vec{v}_1, \ldots, \vec{v}_n$ are mutually orthogonal and all non-zero.

$\vec{v}_1, \ldots, \vec{v}_n$ are a basis of $\vec{V}$

We have proven!

Thm: If $(\, , \,)$ is an inner product on $\mathbb{R}^n$ or any n-dimension $\mathbb{R}$-vector space, $\bar{V}$, and $\mathcal{L}: \bar{V} \to \bar{V}$ s.t.

$$(\mathcal{L}\vec{u}, \vec{v})_{\bar{V}} = (\vec{u}, \mathcal{L}\vec{v})_{\bar{V}}$$

for all $\vec{u}, \vec{v}$, i.e. $\mathcal{L}$ is Self-adjoint, then $\mathcal{L}$ has an orthonormal eigenbasis

$$\left(\begin{array}{l} \text{replace } \vec{v}_i = \vec{v}_i / \|\vec{v}_i\| \\ \text{here } \|\vec{v}\| = \sqrt{(\vec{v}, \vec{v})} \end{array} \right).$$

Question!

Any continuous function on $\vec{x}$

s.t.

$$x_1^2 + x_2^2 + \dots + x_n^2 = 1 \quad (*)$$

has a maximum, why ~~on~~

$\vec{x} \in \{\vec{x} \mid x_1^2 + \dots + x_n^2 = 1\}$ and

$$\left\{ \vec{x} \;\middle|\; \begin{array}{l} (\vec{x}, \vec{v}_1) = 0 \\ (\vec{x}, \vec{v}_2) = 0 \end{array} \right\}$$

Closed! $a \leq x \leq b$ closed condition

$a < x < b$ not

Singular-Value Decomposition:

If $A = (a_{ij}) \in \mathcal{M}_{m,n}(\mathbb{R})$

let

$$\|A\|_{Frob} = \sqrt{\sum_{i,j} |a_{ij}|^2}$$

Same as viewing $A \in \mathbb{R}^{m \times n}$

view as $\mathbb{R}^{mn}$

take usual Euclidean

norm

$$\|A\|_{Frob}^2 = \sum_{i,j} |a_{ij}|^2$$

for computations

$$\sum_{i,j} |a_{ij}|^2 = Tr(AA^T)$$

$$= Tr(A^T A)$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix}$$

$$= \begin{bmatrix} & \\ & \end{bmatrix}$$

$a_{11}^2 + a_{12}^2 + a_{13}^2$ $\qquad$ $a_{21}^2 + a_{22}^2 + a_{23}^2$

$$\mathrm{Tr}(\quad) = \sum_{i,j} a_{ij}^2$$

$=$

Consider

"rank" = dim(image)

$$f(\vec{u}, \vec{v}) = \left\| A - \overbrace{\vec{u}\,\vec{v}^T}^{\text{rank 1 matrix}} \right\|_{\text{frob}}^2$$

$m \times n$

$\vec{u} \in \mathbb{R}^m, \ \vec{v} \in \mathbb{R}^n$

$\vec{u} \hookrightarrow \mathbb{R}^{m \times 1}, \quad \vec{v}^T \hookrightarrow \mathbb{R}^{1 \times n}$

$\vec{u}\,\vec{v}^T \in \mathbb{R}^{m \times n}$

Consider

$$\min_{\substack{\vec{u} \in \mathbb{R}^m \\ \vec{v} \in \mathbb{R}^n}} f(\vec{u}, \vec{v})$$

why?

Imagine that this minimum is attained at $\vec{u} = \vec{u}^*$, $\vec{v} = \vec{v}^*$.

Now: Variational principle $\Rightarrow$ SVD

4 minute break

We consider

$$g(\varepsilon) = f(\vec{u}^*, \underbrace{\vec{v}^* + \varepsilon \vec{w}}_{\text{variation}})$$

we have $g'(0) = 0$

$$C_0 + \varepsilon C_1 + \varepsilon^2 C_2 + \ldots$$

$$C_1 = 0$$

$$g(\varepsilon) = \left\| A - \vec{u}^* (\vec{v}^* + \varepsilon \vec{w})^T \right\|_F^2$$

$$\text{Trace}\left(\left(A - \vec{u}^{*} \vec{v}^{*T} - \vec{u}^{*} \vec{\omega}^{T} \varepsilon\right) \cdot \left(A - \vec{u}^{*} (\vec{v}^{*})^{T} - \vec{u}^{*} (\vec{\omega})^{T} \varepsilon \right)^{T} \right)$$

$$= \text{Trace}\left(\text{mess}_0 + \underbrace{\varepsilon\, \text{mess}_1} + \varepsilon^2 \text{mess}_2 \right)$$

$$\text{mess}_1 = \vec{u}^{*} \vec{\omega}^{T} \left(A - \vec{u}^{*} \vec{v}^{*T}\right)^{T}$$

$$+ \left(A - \vec{u}^{*} \vec{v}^{*T}\right) \left(\vec{u}^{*} \vec{\omega}^{T}\right)^{T}$$

$\text{Trace}(\text{mess}_1)$ must be 0

$$\text{Trace}(B) = \text{Trace}(B^T)$$

$$\Rightarrow$$

$$\text{Trace}(\text{mess}) = 2\,\underbrace{\text{Trace}\left(\left(A - \underbrace{\vec{u}^* \vec{v}^{*T}\right)\vec{w}\,\vec{u}^{*T}}\right)}_{\text{must be } 0}$$

$$\text{Tr}\left(A\vec{w}\,\vec{u}^{*T} - (\vec{v}^* \cdot \vec{w})\,\vec{u}^* \vec{u}^T\right)$$

$$=$$

$$\text{Tr}(CD) = \text{Tr}(DC)$$

$$\operatorname{Trace}\Big(\underbrace{\vec{u}^{*T} A \vec{w}} - (\vec{v}^{*} \cdot \vec{w}) (\vec{u}^{*} \cdot \vec{u}^{*}) \Big)$$

$$\boxed{= \vec{u}^{*T} A \vec{w} - (\vec{v}^{*} \cdot \vec{w}) (\vec{u}^{*} \cdot \vec{u}^{*})}$$

$$\boxed{= 0} \text{ for all } \underline{\vec{w} \in \mathbb{R}^n}.$$

Class ends . . .

Lemma 4.11

Mixing times:

$$d(t+s) \le \bar{d}(s)\,\bar{d}(t)$$

$$d(t) := \max_{\vec{u},\vec{v}\ \text{stochastic}} \left\| \vec{u}P^t - \vec{v}P^t \right\|_{TV}$$

Slick Pf: via coupling.

Pf: optimal coupling from

Prop 4.7

$$\|\mu - \nu\|_{TV} = \inf_{X,Y} \Big\{ \mathrm{Prob}[X \neq Y] \;\Big|\; \overbrace{(X,Y) \text{ is a coupling of } \mu \text{ and } \nu}^{\text{"optimal coupling"}} \Big\}$$

§ 4.2 Coupling & TV distance

Sometimes easier

$$d(t) = \max_{\mu} \|\mu P^t - \pi P^t\|_{TV}$$

$$d(t) \le \bar{d}(t) \le 2d(t)$$

§ 4.2, 4.4, 4.5

Then prove good upper bounds on mixing times in

§ 5.3 Random walk on $\mathbb{B}^n$

" " cycle, torus

Kronecker Prod: 3.18

$$A = (a_{ij}) \in M_{m_1, n_1}(\mathbb{R})$$

$$B = (b_{k\ell}) \in M_{m_2 n_2}(\mathbb{R})$$

$$\underbrace{A \otimes B} \in M_{\underbrace{m_1 m_2}, n, n_2}(\mathbb{R})$$

$$\left(C \ldots \right)$$

here you aren't choosing to flatten

$$C_{\underbrace{(i,k)}, \underbrace{(j,l)}} = a_{ij} b_{kl}$$

using B with blocks of dims of A

$$\left(\quad \right)$$

$$\ldots \quad A \quad - \; - \quad - \quad \quad -$$

$$- \quad B$$

$$\begin{array}{cc} a_{11} & a_{12} \\ \downarrow & \downarrow \end{array} \qquad\qquad \begin{array}{cc} b_{11} & b_{12} \\ \downarrow & \downarrow \end{array}$$

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \otimes \begin{bmatrix} 7 & 8 \\ 9 & 10 \end{bmatrix}$$

$$\left[\begin{array}{c|c} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} 7 & \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} 8 \\ \hline \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} 9 & \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} 10 \end{array}\right]$$

block stable

1st mat is inner

$$\begin{array}{cccc} a_{11}b_{11} & a_{12}b_{11} & a_{11}b_{12} & a_{12}b_{12} \end{array}$$

$$= \left[\begin{array}{cc|cc} 7 & 14 & 8 & 16 \\ 21 & 28 & 24 & 32 \\ 35 & 42 & - & - \\ \hline & & & \end{array}\right]$$

or

$$\begin{bmatrix} 1\begin{bmatrix}7 & 8\\ 9 & 10\end{bmatrix} & 2\begin{bmatrix}7 & 8\\ 9 & 10\end{bmatrix} \\ 3\begin{bmatrix}7 & 8\\ 9 & 10\end{bmatrix} & 4\begin{bmatrix}7 & 8\\ 9 & 10\end{bmatrix} \\ 5\begin{bmatrix}7 & 8\\ 9 & 10\end{bmatrix} & 6\begin{bmatrix}7 & 8\\ 9 & 10\end{bmatrix} \end{bmatrix}$$

$$= \begin{array}{cccc} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ \multicolumn{4}{c}{\left[\begin{array}{cc|cc} 7 & 8 & 14 & 16 \\ \vdots & & & \\ \hline & & & \\ \hline & & & \end{array}\right]} \end{array}$$

block style with 2nd mat as inner

$A_G = a_{ij} = \#$ edges $v_i \to v_j$

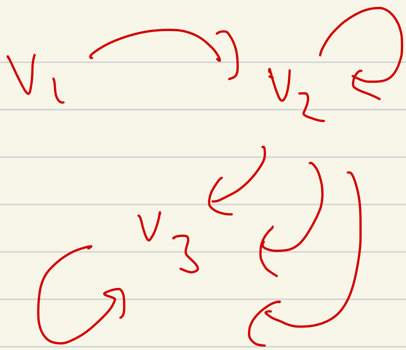
$A_H = b_{kl} = \#$ edges $w_k \to w_l$

$V_G = \{v_1, \dots, v_n\}$

$V_H = \{w_1, \dots, w_m\}$

$(A_G \otimes A_H)_{\underbrace{(v_i, w_k)}_{\text{elt of } V_G \times V_H} \to (v_j, w_l)}$

$= a_{ij} b_{kl} = \begin{pmatrix} \# \text{ edges} \\ v_i \to v_j \end{pmatrix} \cdot \begin{pmatrix} \# \text{ edges} \\ w_k \to w_l \end{pmatrix}$



$w_1 \rightrightarrows w_7$

$(A_G)_{23} = 3$

What is a good definition of $G \otimes_{\text{tensor graphs}} H$ s.t.

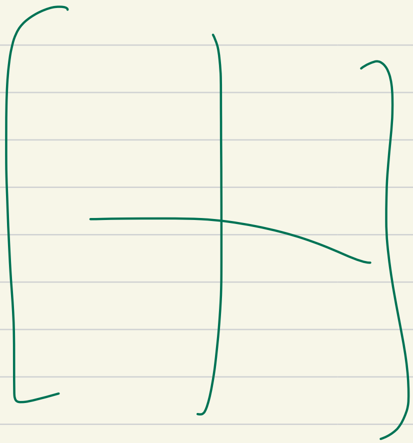
$$\underbrace{A_{\left(G \otimes_{\text{tens prod for graphs}} H\right)}}_{\text{how to define}} = A_G \otimes_{\text{tens prod as matrices}} A_H$$

Defined $G \times H$ simply to $\mathbb{B}^n = (\mathbb{B}^1)^n$

$$H = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \frac{1}{\sqrt{2}}$$

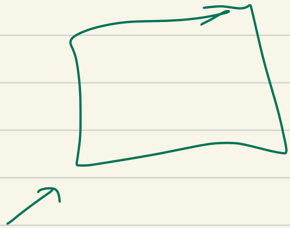
$$H \otimes H$$

same



$$H_1 \otimes H_2$$

fund eigen value of



Square drum

=

sum

fund eig

fund eigen

Separation of variables

$$\Rightarrow \quad u(x,y) = U_1(x) V_1(y) + U_2(x) V_2(y) + \dots$$

functions of x, y $= f(x) g(y)$

Kronecker Prod