March 16, 2021 CPSC 531F

- Today:

Variational Methods:

1. Rayleigh quotients to diagonalize self-adjoint operators

2. SVD (singular-value decomposition)

3. Calculus of variations and other variational methods

Not part of course, "∞-dimensional vector spaces"

Our context!

- $\mathbb{R}^n$, $(\vec{u},\vec{v}) = u_1v_1 + u_2v_2 + \cdots + u_nv_n$

- More generally

$$(\vec{u},\vec{v})_{\vec{w}} = u_1v_1w_1 + \cdots + u_nv_nw_n$$

$\vec{w}$ - weighted dot product

this is mostly enough

More generally yet,

enough $\rightarrow$ any inner product $(\,,\,)$ on $\mathbb{R}^n$

- More generally yet

any inner product $(\,,\,)$ on finite dim vector space

- All this works in inner products $\langle\,,\,\rangle$

over $\mathbb{C}$.

=

Weighted inner product!

$$L : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$

$$\vec{v} \longmapsto A\vec{v}, \quad A \in M_3(\mathbb{R})$$

If $\mathbb{R}^3$ = usual space in which we live,

$$\vec{v} = \begin{bmatrix} 3\text{ cm} \\ 4\text{ cm} \\ 2\text{ cm} \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} 3\text{ ft} \\ 4\text{ ft} \\ 5\text{ ft} \end{bmatrix}$$

but imagine

$$\vec{v} = \begin{pmatrix} 3\text{ cm} \\ 2\text{ km} \\ 5\text{ ft} \end{pmatrix} \ldots$$

For use, weighted inner products

come from
$$\left.\begin{array}{l}(\ ,\ )_{\vec{\pi}} \\ \\ (\ ,\ )_{1/\vec{\pi}}\end{array}\right\} \vec{\pi} \text{ is stationary dist of MC}$$

=

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 6 \\ 3 & 6 & 9 \end{bmatrix},$$

$$\mathcal{L}\begin{bmatrix} 3\text{ cm} \\ 2\text{ km} \\ 5\text{ ft} \end{bmatrix} \overset{?}{=} \begin{bmatrix} ?\text{ cm} \\ ?\text{ light years} \\ ?\ \mu\text{m} \end{bmatrix}$$

$$A = \begin{bmatrix} 1 \ \mathrm{cm}/\mathrm{cm} & 2 \ \mathrm{cm}/\mathrm{cm} & \cdots \\ \vdots & \vdots & \end{bmatrix} \cdots$$

$$= \begin{bmatrix} & 2 \ \mathrm{cm}/\mathrm{cm} & \\ & & \\ & & \end{bmatrix}$$

$$2 \ \mathrm{cm}/\mathrm{cm} = \frac{.02}{1000} \ \mathrm{cm}/\mathrm{km}$$

$$\mathcal{L}\begin{bmatrix} 3\ \text{cm} \\ 2\ \text{km} \\ 5\ \text{ft} \end{bmatrix} = \begin{bmatrix} ?\ \text{cm} \\ ?\ \text{km} \\ ?\ \text{ft} \end{bmatrix}$$

$$\begin{bmatrix} d_1\ \mu\text{m}/\text{cm} & 0 & 0 \\ 0 & d_2\ \mu\text{m}/\text{km} & 0 \\ 0 & 0 & d_3\ \mu\text{m}/\text{ft} \end{bmatrix} \begin{bmatrix} 3\ \text{cm} \\ 2\ \text{km} \\ 5\ \text{ft} \end{bmatrix}$$

$\xrightarrow{\text{then } A}$ $\xrightarrow{\text{then}}$ invert

Simples! Inner Product!

$$\left(\begin{bmatrix} 3\,\text{cm} \\ 2\,\text{km} \\ 5\,\text{ft} \end{bmatrix}, \begin{bmatrix} -2\,\text{cm} \\ 4\,\text{km} \\ 19\,\text{ft} \end{bmatrix} \right)$$

= gives everything in same fixed units

$\rightsquigarrow$ weighted dot product

$$\begin{bmatrix} x_1 \; m \\ x_2 \; km \end{bmatrix} \cdot \begin{pmatrix} y_1 \; m \\ y_2 \; km \end{pmatrix} = \left(x_1 y_1 + x_2 y_2 \, 10^6 \right) m$$

weights

① $\mathcal{L}: \mathbb{R}^n \to \mathbb{R}^n$, $\underbrace{(\ ,\ )}_{\text{inner product}}$

Assume $\mathcal{L}$ is self-adjoint:

$\forall \vec{x}, \vec{y} \in \mathbb{R}^n \qquad (\mathcal{L}\vec{x}, \vec{y}) = (\vec{x}, \mathcal{L}\vec{y})$

(intuition comes from special case:

$(\vec{x}, \vec{y}) = \vec{x} \cdot \vec{y}$, $\mathcal{L}(\vec{v}) \mapsto A\vec{v}$

with $A = A^T$ symmetric)

Thm: Assuming the above, $\mathcal{L}$ has an orthonormal basis of eigenvectors $\vec{v}_1, \ldots, \vec{v}_n$, $\mathcal{L}(\vec{v}_i) = \lambda_i \vec{v}_i$, and $\lambda_i \in \mathbb{R}$.

Proof:

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

- defined for $\vec{v} \neq \vec{0}$
- $R_{\mathcal{L}}(\vec{v}) = R_{\mathcal{L}}(\alpha\vec{v})$ for any $\vec{v} \neq \vec{0}$ and $\alpha \neq 0$.

Claim: $R_{\mathcal{L}}(\vec{v})$ has its maximum value somewhere. Idea: suffices to consider

$$S' = \{ \vec{v} \mid \|\vec{v}\|_2 = \sqrt{v_1^2 + \ldots + v_n^2} = 1 \}$$

(the unit sphere). Then using (Math 320) sequential compactness to prove this.

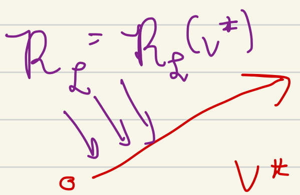
{ $\geq$ 100-150 years ago, did this without proof }

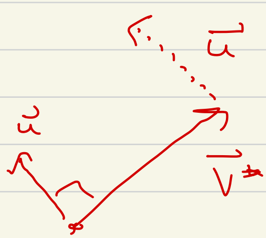
Assume this: we have $\vec{v}^*$ s.t.

$$\max_{\vec{v} \neq 0} R_{\mathcal{L}}(\vec{v}) = R_{\mathcal{L}}(\vec{v}^*).$$

<u>Claim</u>: $\vec{v}^*$ is an eigenvector of $\mathcal{L}$: $\mathcal{L}(\vec{v}^*) = \lambda \vec{v}^*$.

—





Take any $\vec{u} \perp \vec{v}^*$

perpendicular

$$(\vec{u}, \vec{v}^*) = 0$$

Look at

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

"variational"

think of ε near 0:

$$R_\mathcal{L}(\vec{v}^*) \geq R_\mathcal{L}(\vec{v}_\varepsilon) \quad \text{for all } \varepsilon$$

We consider:

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

$$g(\varepsilon) \leq g(0) = R_{\mathcal{L}}(\vec{v}^*) \quad \text{for all } \varepsilon.$$

hence

$$g'(0) = 0$$

now

$$\frac{d}{d\varepsilon} g(\varepsilon) = \frac{d}{d\varepsilon} \left(\frac{(\mathcal{L}\vec{v}_\varepsilon, \vec{v}_\varepsilon)}{(\vec{v}_\varepsilon, \vec{v}_\varepsilon)} \right)$$

$$= \frac{(\vec{v}_\varepsilon, \vec{v}_\varepsilon)\left(\frac{d}{d\varepsilon}(\mathcal{L}\vec{v}_\varepsilon, \vec{v}_\varepsilon)\right) - \text{other way}}{(\vec{v}_\varepsilon, \vec{v}_\varepsilon)^2}$$

$$(\vec{v}_\varepsilon, \vec{v}_\varepsilon) =$$

$$(\vec{v}^* + \varepsilon\vec{u}, \vec{v}^* + \varepsilon\vec{u}) =$$

$$(\vec{v}^*, \vec{v}^*) + 2\varepsilon \underbrace{(\vec{v}^*, \vec{u})}_{0} + \varepsilon^2 (\vec{u}, \vec{u})$$

$$= \underbrace{(\vec{v}^*, \vec{v}^*)}_{\text{non-zero}} + \text{order } \varepsilon^2 \text{ term}$$

$$\frac{1}{(\vec{v}_\varepsilon, \vec{v}_\varepsilon)} = \frac{1}{\underbrace{(\vec{v}^*, \vec{v}^*)}_{\gamma_0} + \text{order}(\varepsilon^2)}$$

$$\gamma_0 + \gamma_1 \varepsilon^2$$

$$= \frac{1}{(\vec{V}^*, \vec{V}^*)} + \text{order}(\varepsilon^2)$$

really $\quad \dfrac{1}{\gamma_0 + \gamma_1 \varepsilon^2} = \dfrac{1}{\gamma_0} \dfrac{1}{1 + (\gamma_1 \varepsilon^2 / \gamma_0)}$

$$= \frac{1}{\gamma_0}\Big(1 - \underbrace{\frac{\gamma_1}{\gamma_0}\varepsilon^2 + \left(\frac{\gamma_1}{\gamma_0}\varepsilon^2\right)^2 + \cdots}_{\substack{\varepsilon \text{ small,} \\ \text{bounded by } C\varepsilon^2}}\Big)$$

$$(\mathcal{L}\vec{V}_\varepsilon, \vec{V}_\varepsilon) =$$

$$\left(\mathcal{L}(\vec{V}^* + \varepsilon \vec{u}), \vec{V}^* + \varepsilon \vec{u}\right)$$

$$= (\mathcal{L}\vec{v}^*, \vec{v}^*) + 2\varepsilon(\mathcal{L}\vec{v}^*, \vec{u}^*) + \underbrace{O(\varepsilon^2)}_{\substack{\text{something bounded} \\ \text{by } \text{Const} \cdot \varepsilon^2}}$$

$$\Longrightarrow$$

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

$$= \frac{1}{(\vec{v}_\varepsilon, \vec{v}_\varepsilon)} \left(\mathcal{L}(\vec{v}_\varepsilon), \vec{v}_\varepsilon \right)$$

$$= \left(\frac{1}{(\vec{v}^*, \vec{v}^*)} + O(\varepsilon^2) \right)$$

$$\left((\mathcal{L} \vec{v}^*, \vec{v}^*) + 2\varepsilon (\mathcal{L} \vec{v}^*, \vec{u}^*) + O(\varepsilon^2) \right)$$

$$= \frac{(\mathcal{L}(\vec{v}^*), \vec{v}^*)}{(\vec{v}^*, \vec{v}^*)} + 2\varepsilon \frac{(\mathcal{L} \vec{v}^*, \vec{u}^*)}{(\vec{v}^*, \vec{v}^*)}$$

$$+ O(\varepsilon^2)$$

In terms of

$g(\varepsilon) = R_{\mathcal{L}}(\vec{v}_\varepsilon)$, $g'(0)$

$$= 2 \frac{(\mathcal{L}(\vec{v}^*), \vec{u})}{\text{something positive}}$$

Hence, for any $\vec{u} \perp \vec{v}^*$

$\left((\vec{u}, \vec{v}^*) = 0 \right)$ we have

$$\mathcal{L}(\vec{v}^*) \perp \vec{u}$$

$$\left(\mathcal{L}(\vec{v}^*), \vec{u} \right) = 0.$$

Hence

$\mathcal{L}(\vec{v}^*)$ is proportional to $\vec{v}^*$

$$\mathcal{L}(\vec{v}^*) = \text{Const} \cdot \vec{v}^* = \lambda \vec{v}^*$$

Why? ...

Next steps:

We can use the same idea to

1) get an entire orthonormal eigenbasis for $\mathcal{L}$

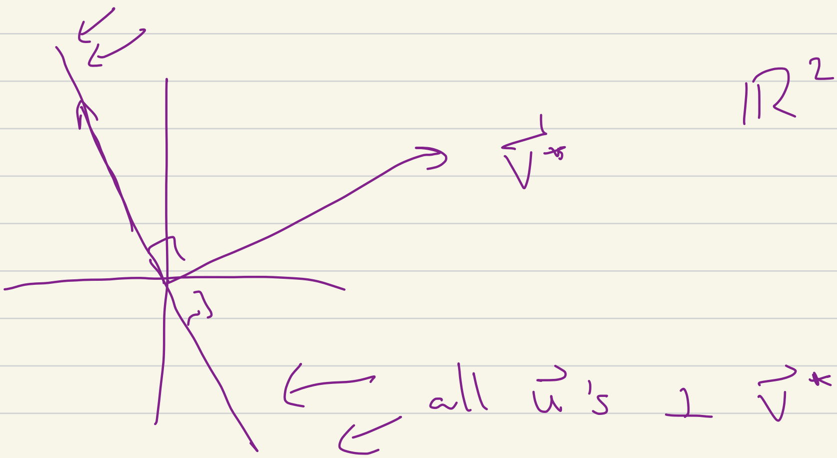
2) Same set of ideas give SVD.

---

4 minute break

Exercise...

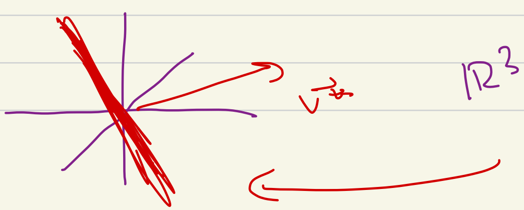
Intuition for dot products



all $\vec{u}$'s $\perp$ $\vec{v}^*$

if $\vec{w} \perp \left(\text{all } \vec{u}\text{'s } \perp \text{ to } \vec{v}^* \right)$

$$\Rightarrow \vec{w} = c \cdot \vec{v}^*$$



stuff $\perp$ $\vec{v}^*$

Everything with $\cdot$ products goes through for $(\, , \,)$ inner products.

Rem! $\quad \mathcal{L}(\vec{u}) = \mu \vec{u} \quad \begin{cases} \vec{u} \neq \vec{0} \\ \vec{u} \in \mathbb{R}^n, \\ \mu \in \mathbb{R} \end{cases}$

then

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

$$= \frac{(\mu \vec{u}, \vec{u})}{(\vec{u}, \vec{u})} = \mu$$

μ eigenvalue $= R_{\mathcal{L}}(\vec{u})$

We know

$$R_{\mathcal{L}}(\vec{v}^*) = \lambda \quad \text{with}$$

$$\mathcal{L}(\vec{v}^*) = \lambda \vec{v}^*$$

and

$$R_{\mathcal{L}}(\vec{v}^*) = \underbrace{\max R_{\mathcal{L}}}_{\substack{\text{at least as big} \\ \text{as a} \\ \text{eigenvalue, } \mu \\ \left(\text{s.t. } \mathcal{L}(\vec{u}) = \mu \vec{u} \text{ for some } \vec{u} \neq 0\right)}}$$

$$\lambda =$$

$\Rightarrow$ λ is the largest eigenvalue of $\mathcal{L}$

Now set

$$\vec{v}_1 = \vec{v}^*, \quad \lambda_1 = \lambda$$

$$= R_{\mathcal{L}}(\vec{v}^*)$$

Consider

$$\max_{\left\{\begin{array}{l}\vec{v} \neq \vec{0} \text{ and} \\ \vec{v} \perp \vec{v}_1\end{array}\right\}} R_{\mathcal{L}}(\vec{v})$$

clever trick

If this max is attained at

$\vec{v}^{**} = \vec{v}_2$, i.e.

$\vec{v}_2 \perp \vec{v}_1$, $\vec{v}_2 \neq 0$, and

$R_{\mathcal{L}}(\vec{v}_2) \geq R_{\mathcal{L}}(\vec{v})$ for any $\vec{v} \perp \vec{v}_1$,

Same variational argument shows that if $\vec{u} \perp$ both $\vec{v}_1, \vec{v}_2$, then $\vec{u} \perp \mathcal{L}(\vec{v}_2)$

$$\Rightarrow \quad \mathcal{L}(\vec{v}_2) = c_1\vec{v}_1 + c_2\vec{v}_2$$

$$\vec{v}_2 \perp \vec{v}_1 \qquad = \qquad \underbrace{c_2\vec{v}_2}$$

we write $\lambda_2\vec{v}_2$

So

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

$$\mathcal{L}(\vec{v}_2) = \lambda_2\vec{v}_2$$

$\vec{v}_2 \perp \vec{v}_1$. Now let $\vec{v}_3$ be the vector $\perp$ $\vec{v}_1, \vec{v}_2$ where

$$\max R_{1}(\vec{v})$$

$\vec{v} \neq \vec{0}$

$\vec{v} \perp \vec{u}_1, \vec{u}_2$

is attained ...

Claim!

SVD comes from almost

the same argument

Class ends

Idea!

$$f(x_1, x_2) = x_1^2 + x_2^2, \qquad x_1, x_2 \in \mathbb{R}$$

where is the min

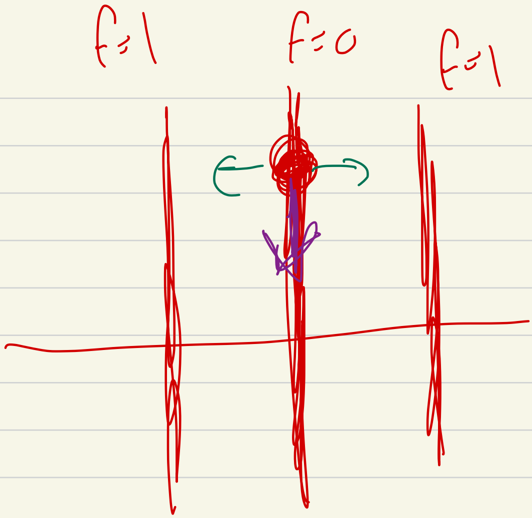
$$\nabla f = (2x_1, 2x_2)$$

$$\nabla f = \vec{0} \Rightarrow x_1 = x_2 = 0$$

$$f(x_1, x_2) = x_1^2$$

$$\nabla f = (2x_1, 0)$$

$$\nabla f = \vec{0} \Rightarrow x_1 = 0, \quad x_2 \text{ anything}$$

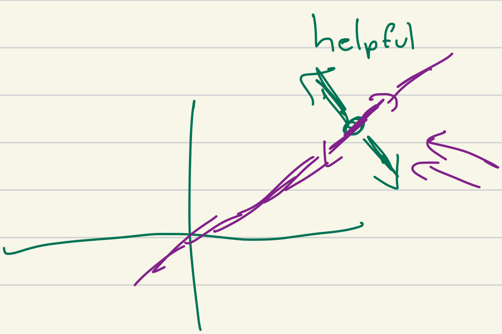


$f(x_1,x_2) = x_1^2$

↓ not helpful

⟷ helpful

$$R_{[2,1]}(x_1,x_2) = \frac{2x_1^2 + x_2^2}{x_1^2 + x_2^2}$$



R = constant

$$R_L(\vec{v}^*)$$

$$= R_L(\alpha \vec{v}^*)$$

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

$$= 3\vec{v}^*$$

$$3\vec{v}^* + \vec{w}$$

$$\Leftarrow \vec{w} \perp \vec{v}^*$$

$$(\vec{v}_\varepsilon, \vec{v}_\varepsilon)$$

$$= (\vec{v}^*, \vec{v}^*) + 2\varepsilon \underbrace{(\vec{v}^*, \vec{u})} + (\vec{u}, \vec{u})\varepsilon^2$$

if $\vec{u} = 3\vec{v}^* + \vec{w}$

$\vec{w} \perp \vec{v}^*$

$$(\vec{v}^*, \vec{u}) = (\vec{v}^*, 3\vec{v}^* + \vec{w})$$

$$= 3(\vec{v}^*, \vec{v}^*) + 0$$

$$\neq 0$$

— EXERCISE —

Solve

Heat Eq $\quad \frac{\partial}{\partial t} u + \Delta u = 0$

Wave Eq $\quad (\frac{\partial}{\partial t})^2 u = \Delta u$

Lapl Eq $\quad \Delta u = 0$, no time

To solve $\quad \vec{u} : \mathbb{R} \to \mathbb{R}^n$, $\quad \vec{u} = \vec{u}(t)$

$$\frac{d}{dt} u = A u \Rightarrow$$

$$\vec{u}(t) = e^{At} \vec{u}(0)$$

want to

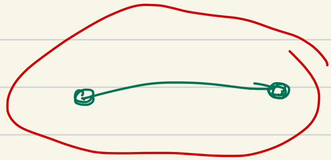
find eigenpairs of A

Want to eigenpairs of Δu

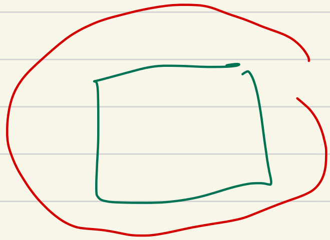
Wave

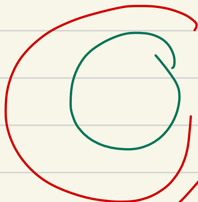
Heat

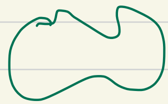
1-dim



2-dim







something worse

Can explicitly find eigenpairs

Can't generally

but eigenpairs of Δ can be found by

$$R_{\Delta} := \frac{(\Delta f, f)}{(f,f)}$$

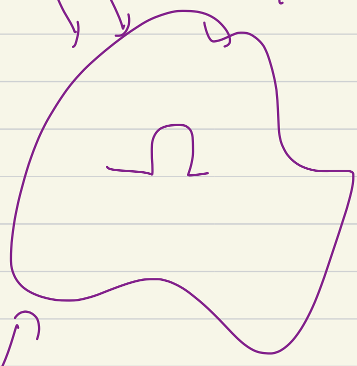
$$= \frac{(\nabla f, \nabla f)}{(f,f)}$$

minimize

$$\frac{(\nabla f, \nabla f)}{(f,f)}$$

over

$\partial\Omega$ boundary of Ω



f s.t. $f\big|_{\partial\Omega} = 0$

gives lowest

eigenfunction
eigenvalue $\Big\}$ of $-\Delta$

wave equation $\Rightarrow$ hearing all the eigenvalues of $-\Delta$