Linear Algebra

15. 12. 2023

Afonso Bandeira

* Please take a look at the typed lecture notes

* The full notes are now available !!

Last Lecture

we were proving

Spectral Theorem

we'll finish the proof today

(quick recall first)

Theorem 6.3.1 [Spectral Theorem]

Any Symmetric matrix $A \in \mathbb{R}^{n \times n}$, $A^T = A$ has a complete set of real eigenvectors and they can be made orthonormal.

(and has n real eigenvalues, up to repetition).

(if $B \in \mathbb{R}^{m \times n}$ not symmetric, $B^T B$ and $B B^T$ are symmetric)

Proof of Spectral Theorem.

$A \in \mathbb{R}^{n \times n}$ sym.

- For any $1 \leq k \leq n$, A has k orthonormal eigenvectors.

$k = n$ gives us our Thm.

Let's use induction.

→ $k = 1$. By Cor 6.3.8 there is a real eigenvector and we can normalize it to have norm 1.

→ Assuming we have $v_1, \dots, v_k$ orthonormal eigenvectors of A (with evalues $\lambda_1, \dots, \lambda_k$) we will show we can "add" one more v_{k+1}.

Let $u_{k+1}, \dots, u_n$ be an orthonormal basis

for the orthogonal complement of the span of
$v_1, \dots, v_k$.

$v_1, \dots, v_k, u_{k+1}, \dots, u_n$ is an orthonormal basis of $\mathbb{R}^n$

V $n \times n$ matrix $V = \begin{bmatrix} | & & | & | & & | \\ v_1 & \cdots & v_k & u_{k+1} & \cdots & u_n \\ | & & | & | & & | \end{bmatrix}$

$V^T V = I$, $V V^T = I$.

$$B = V^T A V = \begin{bmatrix} — & v_1^T & — \\ & \vdots & \\ — & v_k^T & — \\ — & u_{k+1}^T & — \\ & \vdots & \\ — & u_n^T & — \end{bmatrix} \begin{bmatrix} | & | & & | & | & \\ Av_1 & Av_2 & \cdots & Av_k & Au_{k+1} & \cdots \\ | & | & & | & | & \end{bmatrix}$$

$$= \begin{bmatrix} — & v_1^T & — \\ & \vdots & \\ — & v_k^T & — \\ — & u_{k+1}^T & — \\ & \vdots & \\ — & u_n^T & — \end{bmatrix} \begin{bmatrix} | & & | & | & \\ \lambda_1 v_1 & \cdots & \lambda_k v_k & Au_{k+1} & \cdots \\ | & & | & | & \end{bmatrix} = \begin{bmatrix} \text{Nice} & \text{Nice} \\ \text{Nice} & \text{Terrible} \end{bmatrix}$$

Start of 15.12.2023

$$\begin{array}{c|c} & \begin{array}{ccc} & j & \end{array} \\ \begin{array}{c} (k) \\ i \\ \\ (n-k) \\ \ell \end{array} = & \left[\begin{array}{ccc} & \square & \\ & & \\ & & \\ & \square & \\ & & \end{array}\right] \end{array}$$

for $i,j \le k$

$\square = v_i^T \lambda_j v_j = \lambda_j \delta_{ij}$

$j \le k, \ \ell > k$

$\square = v_\ell^T \lambda_j v_j = 0$

$$= \left[\begin{array}{c:c} \Lambda_k & 0_{k \times (n-k)} \\ \hdashline 0_{(n-k) \times k} & C \end{array}\right]$$

Λ_k is $k \times k$ diagonal with $(\Lambda_k)_{ii} = \lambda_i$

C is $(n-k) \times (n-k)$

$C \in \mathbb{R}^{(n-k) \times (n-k)}$

$C^T = C$.

$$B^T = (VAV^T)^T = VA^TV^T = VAV^T = B$$

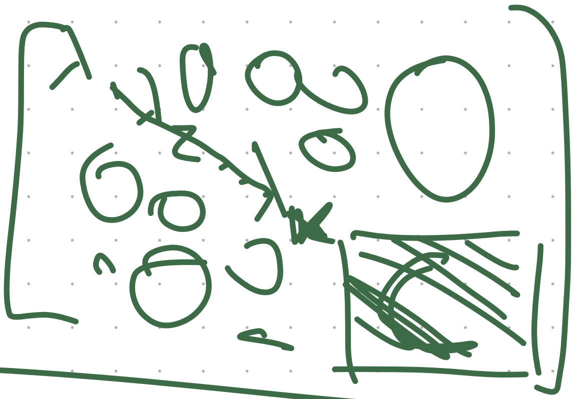
by Cor. 6.3.8 C has a real eigenvalue and eigenvector.

$$Cy = \lambda_{k+1} y$$

$y \in \mathbb{R}^{n-k}$

$y \neq 0$.

$$w = \begin{bmatrix} 0_k \\ y \end{bmatrix} \begin{matrix} \}\, k \text{ zeros} \\ \}\, n-k \text{ entries of } y. \end{matrix}$$

$B =$ 

$$w \in \mathbb{R}^n \qquad w_i := \begin{cases} 0 & \text{if } i \leq k \\ y_{i-k} & \text{if } i > k \end{cases}$$

$$Bw = \underbrace{\left[\begin{array}{c|c} A_k & 0 \\ \hline 0 & C \end{array}\right]}_{B} \underbrace{\left[\begin{array}{c} 0 \\ \hline y \end{array}\right]}_{w} = \left[\begin{array}{c} 0_k \\ \hline Cy \end{array}\right] = \left[\begin{array}{c} 0_k \\ \hline \lambda_{k+1} y \end{array}\right]$$

$$= \lambda_{k+1} \begin{bmatrix} 0 \\ \cdots \\ y \end{bmatrix} = \lambda_{k+1} w.$$

$$B = V^T A V \iff A = V B V^T$$

$$V^T A V w = \lambda_{k+1} w$$

$$\Rightarrow \cancel{V V^T} A V w = \lambda_{k+1} V w$$

I

$$v_{k+1} := V w.$$

set $v_{k+1} := V w$

$$A v_{k+1} = \lambda_{k+1} v_{k+1}$$

$\lambda_{k+1} \in \mathbb{R}$, $v_{k+1} \in \mathbb{R}^n$ and they are e-value & e-vector of A,

we still need to show that

$$v_i^T v_{k+1} = 0 \quad \forall i \leq k.$$

(we can always normalize later)

$$v_i^T v_{k+1} = \left(V^T v_{k+1}\right)_i = w_i = 0 \quad \forall i \leq k$$

QED

Side Note: If $\lambda \in \mathbb{R}$ is an eigenvalue of A the

$N(A - \lambda I)$ has dim ≥ 1.

so there must exist $v \in \mathbb{R}^n \setminus \{0\}$ such that $(A - \lambda I)v = 0$

$$Av = \lambda v.$$

Prop: (Rayleigh quotient)

Given $A \in \mathbb{R}^{n \times n}$ symmetric the Rayleigh quotient is give by

$$R(x) = \frac{x^T A x}{x^T x} \quad \text{for } x \in \mathbb{R}^n \setminus \{0\}$$

$$R : \mathbb{R}^n \setminus \{0\} \longrightarrow \mathbb{R}$$

$$x^T A x = \sum_{i,j=1}^{n} A_{ij} x_i x_j$$

R attains its maximum at

$$R(v_{max}) = \lambda_{max}$$

and its minimum at

$$R(v_{min}) = \lambda_{min}.$$

where λ_{max} is the largest e-value of A and v_{max} it's corresponding e-vector

and λ_{min} is the smallest ....

v_{min} ....

Proof:

$$R(v_{max}) = \frac{v_{max}^T A v_{max}}{v_{max}^T v_{max}}$$

$$= \frac{v_{max}^T \lambda_{max} v_{max}}{v_{max}^T v_{max}} = \lambda_{max}.$$

and similarly $R(v_{min}) = \lambda_{min}$.

Now we need to show

$$\lambda_{min} \leq R(x) \leq \lambda_{max}$$

$$A = \sum_{i=1}^{n} \lambda_i v_i v_i^T \quad \left(\begin{array}{l} v_i\text{'s e-vectors of } A \\ \lambda_i\text{'s eigenvalues} \end{array} \right)$$

$$R(x) = \frac{x^T \left(\sum_{i=1}^{n} \lambda_i v_i v_i^T \right) x}{x^T x} =$$

$$= \frac{1}{\|x\|^2} \sum_{i=1}^{n} \lambda_i (x^T v_i)(v_i^T x)$$

$$= \frac{1}{\|x\|^2} \sum_{i=1}^{n} \lambda_i \underbrace{(x^T v_i)^2}_{\geq 0.}$$

$$\lambda_{min} (x^T v_i)^2 \leq \lambda_i (x^T v_i)^2 \leq \lambda_{max} (x^T v_i)^2$$

$$\frac{1}{\|x\|^2}\sum_{i=1}^{n}\lambda_{min}(x^Tv_i)^2 \le R(x) \le \frac{1}{\|x\|^2}\sum_{i=1}^{n}\lambda_{max}(x^Tv_i)^2$$

$$\sum_{i=1}^{n}(x^Tv_i)^2 = \sum_{i=1}^{n}(v_i^Tx)^2$$

$$= \|V^Tx\|^2 = \|x\|^2.$$

$$\lambda_{min} \le R(x) \le \lambda_{max}$$

$\square$.

11

Def: A sym matrix $A \in \mathbb{R}^{n\times n}$ is called Positive Semidefinite (PSD) if all eigenvalues of A are non-negative (≥ 0). If they are all positive (>0) then A is Positive Definite (PD).

Ex: Proj matrices are PSD.

Prop: 6.3.12. A is PSD ($A \succeq 0$) (A is $n\times n$) iff $x^TAx \geq 0 \quad \forall x \in \mathbb{R}^n$

(and A is PD if $x^TAx > 0 \quad \forall x \in \mathbb{R}^n \setminus \{0\}$)

proof.: Follows fr Prop. about Rayleigh quotient

If $\lambda < 0$ is e-value of A, $v_{min}^T A v_{min} < 0$.

if $\lambda_{min} \geq 0$ then $R(x) \geq \lambda_{min} \geq 0$

and so

$$\frac{x^T A x}{x^T x} \geq 0 \text{ and } x^T x > 0 \text{ if } x \neq 0.$$

if $x \neq 0$ $x^T A x \geq 0$

if $x = 0$ $x^T A x = 0$.

<u>Fact 6.3.13:</u> If A and B are both PSD then $A+B$ is PSD.

<u>Proof:</u> $x^T(A+B)x = x^T A x + x^T B x \geq 0$. ✓

—||—

Def: Gram Matrix. Given $v_1, \ldots, v_m \in \mathbb{R}^n$

$G \in \mathbb{R}^{m \times m}$ $G_{ij} := v_i^T v_j$

$$V = \begin{bmatrix} v_1 v_2 \cdots v_m \end{bmatrix}, \quad G = V^T V.$$

V is $n \times m$

G is sym.

G is $m \times m$

Remark:

given $v_1, \dots, v_m \in \mathbb{R}^n$ we can also build

$$H = V V^T$$

$$H = \sum_{i=1}^{m} v_i v_i^T$$

(sometimes also called a Gram matrix)

H is sym

H is $n \times n$.

Prop: $A^T A$ and $A A^T$ have the same non-zero eigenvalues.

mini CS Les

$$\begin{array}{cc} & \text{person} \\ \text{feature} & \begin{bmatrix} & A & \end{bmatrix} \end{array}$$

$A^T A$ people x people

$A A^T$ feature x feature

---

eigenquiz game