Lecture 7.1: Consequences of the spectral theorem

Optimization and Computational Linear Algebra for Data Science

Léo Miolane

The Spectral theorem

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a **symmetric** matrix. Then there is a orthonormal basis of $\mathbb{R}^n$ composed of eigenvectors of A.

That means that if A is symmetric, then there exists an orthonormal basis $(v_1, \dots, v_n)$ of $\mathbb{R}^n$ and $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ such that

$$\boxed{Av_i = \lambda_i v_i \qquad \text{for all} \quad i \in \{1, \dots, n\}.}$$

Theorem (Matrix formulation)

Let $A \in \mathbb{R}^{n \times n}$ be a **symmetric** matrix. Then there exists an orthogonal matrix P and a diagonal matrix D of sizes $n \times n$ such that

$$\boxed{A = PDP^{\mathsf{T}}.}$$

Consequences

If

$$\boxed{A = P \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} P^{\mathsf{T}}}$$

for some orthogonal matrix P then:

Consequence #1: $\lambda_1, \ldots, \lambda_n$ are the only eigenvalues of A, and the number of time that an eigenvalue appear on the diagonal equals its multiplicity.

Proof sketch on an example

Consider $n = 3$ and

$$A = P \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix} P^{\mathsf{T}} \qquad \text{where} \qquad P = \begin{pmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{pmatrix}$$

is an orthogonal matrix.

$$\left|\begin{array}{l} A v_1 = PDP^{\mathsf{T}} v_1 = PD e_1 = P(3e_1) = 3v_1 \\ A v_2 = 3 v_2 \\ A v_3 = -1\, v_3 \end{array}\right.$$

- $\dim E_3(A) \geqslant 2$ and $\dim E_{-1}(A) \geqslant 1$
- $\boxed{\dim E_3(A) + \dim E_{-1}(A) \leqslant 3}$
- therefore $\left|\begin{array}{l} \dim E_3(A) = 2 \\ \dim E_{-1}(A) = 1 \end{array}\right.$ and there is no other eigenvalues.

Consequences

If

$$A = P \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} P^\mathsf{T}$$

for some orthogonal matrix P then:

Consequence #2: The rank of A equals to the number of non-zero λ_i's on the diagonal:

$$\boxed{\operatorname{rank}(A) = \#\{i \,|\, \lambda_i \neq 0\}.}$$

Proof

Using HW 3 we have

$$\mathrm{rank}(A) = \mathrm{rank}(PDP^T) \quad \text{(because } P^T \text{ is invertible)}$$

$$= \mathrm{rank}(PD) \quad \text{(—— } P \text{ ——)}$$

$$= \mathrm{rank}\left(\begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}\right)$$

$$= \#\{ i \mid \lambda_i \neq 0 \}$$

Consequences

If

$$A = P \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} P^{\mathsf{T}}$$

for some orthogonal matrix P then:

Consequence #3: A is invertible if and only if $\boxed{\lambda_i \neq 0 \text{ for all } i}$. In such case

$$\boxed{A^{-1} = P \begin{pmatrix} 1/\lambda_1 & 0 & \cdots & 0 \\ 0 & 1/\lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1/\lambda_n \end{pmatrix} P^{\mathsf{T}}}$$

Proof

A is invertible $\iff$ $\operatorname{rank}(A) = n$ (Consequence 2)

$\iff$ $\lambda_i \neq 0$ for all i

Assume $\lambda_i \neq 0$ for all i then:

$$\overbrace{P \operatorname{Diag}(\lambda_1 \dots \lambda_n) P^T}^{A} \underbrace{P}_{= Id} \operatorname{Diag}(\lambda_1^{-1}, \dots, \lambda_n^{-1}) P^T = A^{-1}$$

$$= P \operatorname{Diag}(1, 1, \dots, 1) P^T$$

$$= Id.$$

Consequences

If

$$A = P \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} P^{\mathsf{T}}$$

for some orthogonal matrix P then:

Consequence #4: $\boxed{\operatorname{Tr}(A) = \lambda_1 + \cdots + \lambda_n.}$

$$\operatorname{Tr}(A) = \operatorname{Tr}(PDP^T) = \operatorname{Tr}(\underbrace{P^T P}_{=Id} D)$$

HW 3,4?

$$= \operatorname{Tr}(D) = \lambda_1 + \dots + \lambda_n$$