The main theorem of diagonalization

Theorem 1

The following conditions for A are equivalent

- A is diagonalizable
- A satisfies a polynomial of simple roots
- A is a linear combination of projection matrices $P_1, ..., P_k$ such that $P_1 + \cdots + P_k = I$.

We prove this theorem by proving

Condition 1 $\implies$ Condition 2 $\implies$ Condition 3 $\implies$ Condition 1

The main theorem of diagonalization

Condition 1 $\implies$ Condition 2.

Suppose A is diagonalizable, we can write

$$A = P^{-1} \begin{pmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \ddots & & \\ & & & \lambda_{n-1} & \\ & & & & \lambda_n \end{pmatrix} P.$$

Note that some of the diagonal entries might be repeated, therefore, we collect them into a set

$$S = \{\lambda_1\} \cup \{\lambda_2\} \cup \cdots \cup \{\lambda_n\}$$

therefore we compose polynomial

$$f(t) = \prod_{\mu_i \in S} (t - \mu_i)$$

The main theorem of diagonalization

We have

$$f(A) = P^{-1} \begin{pmatrix} f(\lambda_1) & & & & \\ & f(\lambda_2) & & & \\ & & \ddots & & \\ & & & f(\lambda_{n-1}) & \\ & & & & \lambda_n \end{pmatrix} P = 0$$

The main theorem of diagonalization

Condition 2 $\implies$ Condition 3.

Now if A satisfies

$$(A-\lambda_1)(A-\lambda_2)\cdots(A-\lambda_n)=0$$

for **distinct** $\lambda_1,\cdots,\lambda_n$, then we may construct interpolation polynomials $f_{\lambda_1},\cdots,f_{\lambda_n}$ with the property that $f_{\lambda_j}(\lambda_i)=0$ for any $i\neq j$ and $f_{\lambda_i}(\lambda_i)=1$. Using interpolation theorem, we interprating the constant function $g(t)=1$, we have

$$1=Q(t)(t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_n)+f_{\lambda_1}(t)+f_{\lambda_2}(t)+\cdots+f_{\lambda_k}(t)$$

letting $g(t)=t$, we have

$$t=Q(t)(t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_n)+\lambda_1 f_{\lambda_1}(t)+\lambda_2 f_{\lambda_2}(t)+\cdots+\lambda_k f_{\lambda_k}(t)$$

The main theorem of diagonalization

Plug in $t = A$ to those two polynomials, we denote

$$P_{\lambda_i} := f_{\lambda_i}(A),$$

we know P_{λ_i} are projection matrices, therefore

$$I = P_{\lambda_1} + P_{\lambda_2} + \cdots + P_{\lambda_k}$$

and

$$A = \lambda_1 P_{\lambda_1} + \lambda_2 P_{\lambda_2} + \cdots + \lambda_k P_{\lambda_k}$$

The main theorem of diagonalization

Condition 3 $\implies$ Condition 1.

Recall that by cross filling, we essentially decompose each projection matrices into rank 1 projection matrices

$$P = \vec{v}_1 \vec{w}_1^T + \cdots$$

Therefore, A is also a linear combination of rank-1 proejction matrices

$$A = \mu_1 \vec{v}_1 \vec{w}_1^T + \mu_2 \vec{v}_2 \vec{w}_2^T + \cdots + \mu_n \vec{v}_n \vec{w}_n^T.$$

$$I = \vec{v}_1 \vec{w}_1^T + \vec{v}_2 \vec{w}_2^T + \cdots + \vec{v}_n \vec{w}_n^T.$$

Note that each rank one projection must be trace one. Since the trace of I is n, we know there are **exactly** n many rank-1 projections. Therefore we may write it into matrix form

$$A = \begin{pmatrix} | & | & \cdots & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & \cdots & | \end{pmatrix} \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix} \begin{pmatrix} - & \vec{w}_1^T & - \\ \vdots & \vdots & \vdots \\ - & \vec{w}_n^T & - \end{pmatrix}$$

The main theorem of diagonalization

and

$$\begin{pmatrix} | & | & \cdots & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & \cdots & | \end{pmatrix} \begin{pmatrix} - & \vec{w}_1^T & - \\ \vdots & \vdots & \vdots \\ - & \vec{w}_n^T & - \end{pmatrix} = I$$

Since the two matrices are square matrices, so let

$$P = \begin{pmatrix} | & | & \cdots & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & \cdots & | \end{pmatrix}$$

Then

$$P^{-1} = \begin{pmatrix} - & \vec{w}_1^T & - \\ \vdots & \vdots & \vdots \\ - & \vec{w}_n^T & - \end{pmatrix}$$

We have $A = P \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix} P^{-1}$, therefore A is diagonalizable.

The main theorem of diagonalization

Let′s think about the following equivalent conditions

- A is diagonalizable
- A satisfies a polynomial of simple roots
- A is a linear combination of projection matrices $P_1, ..., P_k$ such that $P_1 + \cdots + P_k = I$.

Excercise. Suppose $A = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_m P_m$ is a linear combination of matrices such that $P_1 + P_2 + \cdots + P_m$ is a projection, is A diagonalizable? why?

Solution. Denote by $P := P_1 + P_2 + \cdots + P_m$, since P is a projection, so is $I - P$. Let $P_{m+1} = I - P$, we have

$$P_1 + P_2 + \cdots + P_m + P_{m+1} = I$$

and

$$A = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_m P_m + \underbrace{0}_{\lambda_{m+1}} P_{m+1}$$

Therefore A is diagonalizable

The main theorem of diagonalization

Can you find a linear combination of projection matrices such that it is not diagonalizable?

The main theorem of diagonalization

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}}_{\text{NOT DIAGONALIZABLE!}}$$

Excercise. Diagonalize the following matrix

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

The main theorem of diagonalization

Excercise. If

$$A = \lambda_1 P_1 + \cdots + \lambda_m P_m$$

however,

$$P_1 + \cdots + P_m \neq I,$$

but

$$(P_1 + \cdots + P_m)^2 = P_1 + \cdots + P_m$$

is a projection. Is A a diagonalizable matrix ? why?