特征值与特征向量 相似对角化

相似

相似的定义

与相抵标准形类似的，对一个 n 维线性空间 V 上的线性变换$\sigma \in L(V)$，找到一组基{α_1,...,α_n}，
使得 T 在这组基下的矩阵就像相抵标准形那样，有尽可能多的0，即我们有
σ(α_1,...,α_n) = (α_1,...,α_n)A,
我们希望A有尽可能多的0，非零元素的排列规律也尽可能简单.

假设我们已有了一组基，那么最直接的就是引入基变换，利用变换矩阵P，使得
$A' = P^{-1}AP$，那么只需要让 A'尽可能简单。

通过基变换对矩阵的影响：

设线性变换$\sigma \in \mathcal{L}(V,V)$，$B_1 = \{\alpha_1, \ldots, \alpha_n\}$和$B_2 = \{\beta_1, \ldots, \beta_n\}$是线性空间的$V(\mathbf{F})$的两组基，基$B_1$变为基$B_2$的变换矩阵为$C$，如果$\sigma$在基$B_1$下的矩阵为$A$，则$\sigma$关于基$B_2$所对应的矩阵为$C^{-1}AC$.

--

我们可以得到一个合适的定义。

若存在可逆矩阵 P，使得

$$P^{-1}AP = B,$$

则称 A 与 B **相似**，记作 $A \sim B$。

性质

- 相似的两个矩阵必须都是方阵。
- 相似矩阵有相同的特征值、迹、行列式、秩。
- 相似关系是等价关系。（自反、对称、传递）
- 相似必定相抵，但反之不一定成立。
- 对于任意多项式$f(x)$都有$f(A) \sim f(B)$，且若$B = P^{-1}AP$，有$f(B) = P^{-1}f(A)P$. 除此之外还有A, B可逆时，$A^{-1} \sim B^{-1}$，$A^* \sim B^*$

--

- $A_1 \sim B_1$，$A_2 \sim B_2$不一定有$A_1 + A_2 \sim B_1 + B_2$；
  只有当$P^{-1}A_1P = B_1, P^{-1}A_2P = B_2$时（即相同的过渡矩阵$P$）才有$P^{-1}(A_1 + A_2)P = B_1 + B_2$
- 若$A_1 \sim B_1$，$A_2 \sim B_2$，则有

$$\begin{pmatrix} A_1 & O \\ O & A_2 \end{pmatrix} \sim \begin{pmatrix} B_1 & O \\ O & B_2 \end{pmatrix};$$

特征值、特征向量、特征子空间及其求解

特征值与特征向量

$V(\mathbf{F})$上的一个线性变换总是把线性空间上的一些向量映射到另一些向量，那么这个过程当中就会有一些向量的方向并不会改变，而只是简单的伸缩。在探讨相似的过程当中，两个相似的矩阵，处理同样的一些向量时，我们会观察到同样的变化情况。

设σ是线性空间$V(\mathbf{F})$上的一个线性变换，如果存在数$\lambda \in \mathbf{F}$和非零向量$\xi \in V$使得$\sigma(\xi) = \lambda\xi$，则称数λ为σ的一个特征值，并称非零向量ξ为σ属于其特征值λ的特征向量。

同样的，对线性空间F^n上的矩阵A，若存在数$\lambda \in \mathbf{F}$和非零向量$X \in V$，使得$AX = \lambda X$，则称数λ为A的一个特征值，并称非零向量X为A属于其特征值λ的特征向量。

必须注意特征向量为非零向量，否则零向量$\xi = \vec{0}$对任意λ都满足上面定义，从而失去"特征"的含义. 但是特征值可以为0，此时$\sigma(\xi) = \vec{0}$，即全体特征向量的集合就是线性变换的核空间.

特征子空间

对于某一个$\lambda \in \mathbf{F}$，我们将所有满足$\sigma(\xi) = \lambda\xi$的向量构成的集合记为$E(\lambda, \sigma) = \{\xi \mid \sigma(\xi) = \lambda\xi,\ \xi \in V\}$（在去除线性变换不引起歧义的情况下可简写为$V_\lambda$），称为$\sigma$关于其特征值$\lambda$的特征子空间. 显然，这一集合是由零向量和全体$\lambda$对应的特征向量构成的. 我们可以验证$V_\lambda$的确是$V$的子空间：

--

V_λ的维数至少是1，但不一定是1。
当维数不止是1时，一个特征值可能会对应不止一个特征向量。

此时我们称这个特征子空间V_λ的维数为这个特征值λ的几何重数。

特征值与特征向量的计算

设σ是$V(\mathbf{F})$上的线性变换，I为恒等映射，则下述条件等价：

$\lambda \in \mathbf{F}$是σ的特征值；

$\sigma - \lambda I$不是单射；

$\sigma - \lambda I$不是满射；

$\sigma - \lambda I$不可逆.

由此，我们有一个重要的计算方式：

$\lambda \in \mathbf{F}$是σ的特征值等价于$\sigma - \lambda I$不可逆，因此其在V的任意一组基$\alpha_1, ..., \alpha_n$下的矩阵$A - \lambda E$也不可逆（其中A为σ在这组基下的矩阵，E为单位矩阵），这又等价于$|A - \lambda E| = 0$。

设$f(\lambda) = |\lambda E - A|$。由于$f$是在之后的讨论中有核心地位的概念，我们称其为矩阵$A$的特征多项式，其$k$重根称为$k$重特征值（称$k$为代数重数）。

由此，我们得到了一个通过求解特征多项式来求出特征值，进而得到特征向量的办法。

特征多项式的性质

对于$A = (a_{ij}) \in \mathbf{M}_n(\mathbf{F})$，记

$$f(\lambda) = |\lambda E - A| = a_0\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$$

则$a_0 = 1$，$a_1 = -tr(A)$，$a_n = (-1)^n|A|$，且a_k等于所有k级主子式之和乘以$(-1)^k$。

这里我们主要掌握两个特例，即由韦达定理，我们有

$$\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii};$$

$$\prod_{i=1}^{n} \lambda_i = |A|.$$

即特征值按重数求和为矩阵的迹（即矩阵对角线元素之和），特征值按重数求积为矩阵行列式. 这一结论在解决某些问题时有一定作用.

相似矩阵有相同的特征多项式（逆命题不成立），即$A \sim B \implies |\lambda E - A| = |\lambda E - B|$，
从而有相同的迹，行列式，特征值，但特征向量不一定相同。

若A的多项式$f(A) = O$，那么事实上A的特征值必然是$f(\lambda) = 0$的根。

设$A\in\mathbf{F}^{n\times n}$，且存在多项式$f(x)\in\mathbf{F}[x]$使得

$$f(A)=a_mA^m+a_{m-1}A^{m-1}+\cdots+a_1A+a_0E=O,$$

则对于任意的$\lambda\in\mathbf{F}$，$f(\lambda)\neq 0\implies A-\lambda E_n$可逆.

等价地，对于任意的$\lambda\in\mathbf{F}$，若λ为A的特征值，则$f(\lambda)=0$

有关特征值的性质：

关于特征值，我们有如下基本性质：

设λ是线性空间$V(\mathbf{F})$上的线性变换σ的特征值，ξ是σ属于λ的特征向量，则

- $k\lambda$是$k\sigma$的特征值，λ^m是σ^m的特征值，且ξ仍是相应特征向量；
- 若$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$是$\mathbf{F}$上的多项式，则$f(\sigma)(\xi)=f(\lambda)\xi$；

设λ是n阶矩阵A的特征值，A可逆，则λ^{-1}是A^{-1}的特征值，$|A|\lambda^{-1}$是A的伴随矩阵A^*的特征值，且特征向量不变.

设A为n阶矩阵，则A与A^{T}有相同的特征值（含重数）.

有关特征向量的性质

设V是有限维的，$\sigma\in\mathcal{L}(V)$且$\lambda\in\mathbf{F}$，则

- \sigma$的不同特征值对应的特征向量线性无关；
- σ的不同特征值对应的特征子空间的和为直和；
- σ最多有$\dim V$个不同的特征值.

进一步的，我们有：

n维线性空间$V(\mathbf{F})$的线性变换σ的每个特征值λ_0的重数（代数重数）大于等于其特征子空间V_{λ_0}的维数（几何重数）。

相似对角化

对角化的定义和基本性质

设$\sigma\in\mathcal{L}(V)$，如果存在V的一组基使得σ在这组基下的矩阵是对角矩阵，则称σ可对角化.

上述定理的证明蕴含着一个重要的结论：如果一个线性变换可对角化，那么对应于对角矩阵的基中每一个向量必然是特征向量，因此我们很容易想到可对角化的下一个等价条件就是线性变换有一组由特征向量构成的基. 当然这只是一个观察，基于这一观察我们可以进一步扩展，得到下面这一关于线性变换可对角化等价条件的核心定理：

设V是数域$\mathbf{F}$上的n维线性空间，σ是V上的线性变换，$\lambda_1,\lambda_2,\ldots,\lambda_s\in\mathbf{F}$是$\sigma$的所有互异特征值，则以下条件等价：

- σ可对角化；
- σ有n个线性无关的特征向量，它们构成V的一组基；
- $V有在\sigma下不变的一维子空间U_1,\ldots,U_n， 使得V=U_1\oplus\cdots\oplus U_n$.
- $V=V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_s}$；
- $n=\dim V_{\lambda_1}+\dim V_{\lambda_2}+\cdots+\dim V_{\lambda_s}$
- σ每个特征值的代数重数等于几何重数.

我们有一个显然的推论如下：

若n维空间上的线性变换σ有n个不同的特征值，则σ可对角化. 反之，σ可对角化不一定有n个特征值.

矩阵的相似对角化

设$A\in\mathbf{F}^{n\times n}$，如果存在可逆矩阵$P$使得$P^{-1}AP$是对角矩阵，则称$A$可对角化（等价于$A$相似于对角矩阵）

可对角化的线性映射的矩阵都是可对角化的。

矩阵可对角化的等价条件：我们可以将以上线性变换可对角化的结果都推广到矩阵的可对角化上：

设A是数域$\mathbf{F}$上的n阶矩阵，$\lambda_1,\lambda_2,...,\lambda_s\in\mathbf{F}$是$A$的所有互异特征值，则以下条件等价：

- A可对角化；
- A有n个线性无关的特征向量，它们构成$\mathbf{F}^n$的一组基；
- $n=\dim V_{\lambda_1}+\dim V_{\lambda_2}+\cdots+\dim V_{\lambda_s}$
- A每个特征值的代数重数等于几何重数.

同理可得：

若n阶矩阵A有n个不同的特征值，则A可对角化. 反之，A可对角化不一定有n个特征值.

一般的对角化计算

一个简明的步骤：

- 先任意写出σ在一组基B下的矩阵A，当然为了计算方便一般选取自然基；
- 利用特征多项式$f(\lambda)=|\lambda E-A|=0$求出$A$的所有不同特征值；
- 解线性方程组$AX=\lambda X$（实际上就是方程组$(\lambda E-A)X=0$，其中λ是上一步求出的特征值）求出A在不同特征值下的线性无关特征向量；
- 第三步中求得的所有向量就是σ的特征向量在基B下的坐标，根据前面的讨论，σ的特征向量也就是使得σ的矩阵表示为对角矩阵的那组基.
- 当然，如果题目中直接给出求P使得$P^{-1}AP$为对角矩阵，那么我们只需进行2、3两步，并将3中得到的向量按列排列成矩阵P即可.

一些特殊的矩阵

一个矩阵A（或线性变换σ）称为对合矩阵（或对合变换），如果$A^2=E$（或$\sigma^2=I$）；
称为幂等矩阵（或幂等变换），如果$A^2=A$（或$\sigma^2=\sigma$）；称为幂零矩阵（或幂零变换），如果存在自然数k使得$Ak=O$（或$\sigma k=O$）

对幂等矩阵，我们还有一些特别的性质：

命题（幂等矩阵“诱导”出的矩阵仍是幂等矩阵）
设 A 为 n 阶幂等矩阵，I_n 为 n 阶单位矩阵，则

1. 若 B 与 A 相似，则 B 是幂等矩阵；
2. A^T、A^*、I_n-A^T 都是幂等矩阵；
3. 对任何正整数 k，A^k 也是幂等矩阵。

命题（幂等矩阵的行列式）
若 A 为幂等矩阵，则 $\det(A)=0$ 或 $\det(A)=1$。

证明：根据行列式的性质，有 $\det(A)=\det(A^2)=\det^2(A)$，从而 $\det(A)=0$ 或 $\det(A)=1$。

命题（幂等矩阵的特征值）
若 A 为幂等矩阵，则 A 的特征值只能为 0 或 1。

证明：设 $\lambda\in\mathbb{C}$ 是 A 的一个特征值，v 是 A 关于 v 的一个特征向量，则由特征值定义有

$$Av=\lambda v$$

$$A^2v=AAv=\lambda v$$

联立两式，就有 $\lambda^2v=\lambda v$。因为 v 非零，所以 $\lambda=1$ 或 $\lambda=0$。

命题（幂等矩阵的像空间“不变”）
若 A 为幂等矩阵，则 $Ax=x$ 的充要条件是 $x\in\operatorname{im}L_A$。

证明：必要性是显然的，我们证明充分性。
因为 $x\in\operatorname{im}L_A$，所以存在 $w\in\mathbb{F}^n$，使得 $Aw=v$。又 $A^2w=v$，$A^2w=AAw=Av$，所以 $Av=v$。

命题3讨论了幂等矩阵的特征值，由此自然我们会关心，幂等矩阵能否相似对角化？答案是肯定的，我们会在定理1中给出证明。而为了证明定理1，先给出下面两条与线性映射有关的引理。

引理1
设 T 为有限维线性空间 V 上的线性算子，则下面两条性质等价：

1. $\ker T + \operatorname{im} T = V$；
2. $\ker T \cap \operatorname{im} T = \{0\}$。

引理2（幂等矩阵对应的直和分解）
若 A 为幂等矩阵，则 $\ker L_A \oplus \operatorname{im} L_A = V$。

证明：由引理1，我们只需证明 $\ker L_A \cap \operatorname{im} L_A = \{0\}$。
设 $v \in \ker L_A \cap \operatorname{im} L_A$，有 $Av = 0$，依命题4又有 $Av = v$，于是 $v = 0$，即 $\ker L_A \cap \operatorname{im} L_A = \{0\}$。

定理1（幂等矩阵可相似对角化）
设 n 阶幂等矩阵的秩为 $r\ (1 \le r \le n)$，则 A 相似于对角阵 $\begin{pmatrix} I_r & \\ & O \end{pmatrix}$。

根据定理，以下推论是显然的：

推论1
若幂等矩阵 A 的秩为 r，则 $\operatorname{tr}(A) = r$。

推论3
若 A 为 n 阶实对称幂等矩阵，则存在正交矩阵 Q，使得 $Q^{-1}AQ = \begin{pmatrix} I_r & \\ & O \end{pmatrix}$。

习题

直接用LALU上的吧。