

## 2.12 第二单元复习

第二单元的后半部分都是围绕矩阵特征值与特征向量的应用来讲的，其实这也是我们后续研究矩阵的主要方式！

strang教授同样以习题的形式来带大家复习！

### 知识回顾

#### 投影部分

首先我们讨论了正交性，如果两个向量点乘为0，那么他们就正交。而我们有一个矩阵  $Q = [q_1 \ \dots \ q_n]$ ，当所有q都是标准正交基时，我们称Q为正交矩阵！那么此时他有性质  $Q^T Q = I$ 。

然后我们学习了投影，推导了投影矩阵，然后通过投影矩阵去解决了  $Ax=b$  无解的情况，通过最小二乘法去拟合，寻找最优解！

然后我们就学习了Gram-Schmidt 正交化方法，把一组线性无关的向量通过投影矩阵的操作，量投影到另一组向量上，新得到的向量正交，再进行单位化，将基变为标准正交基。

这大概就是投影部分学的内容！

#### 行列式部分

首先介绍了行列式的十个性质，其中最重要的是 1,2,3；而后面的性质都是由前三个性质导出。

然后我们根据性质1，2，3推导出了行列式的展开式，然后我们根据展开式得到了代数余子式，这也是我们计算行列式的简便方法之一。最后得到逆矩阵的求解公式！

#### 特征值部分

首先介绍了特征值与特征向量， $Ax = \lambda x$  方程，以及求特征向量的方程。另外，如果矩阵有 n 个线性无关的特征向量，那么将它们构成矩阵 S，可以用来将矩阵进行对角化处理： $A = S\Lambda S^{-1}$ 。同时，可以使用对角化公式计算矩阵幂的问题。然后引入了许多应用，例如微分方程之类的问题（这样是本门课程中讨论过最难的几个问题之一）。

那么我们开始看题目吧！

### (1)

现在我们有一个向量  $a = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$ ，找到向量 a 所在直线的投影矩阵，然后找到投影矩阵的特征向量与特征值！

首先我们求解投影矩阵的方法就是运用公式  $A(A^T A)^{-1} A^T$ ，代入得  $a(a^T a)^{-1} a^T$ ，由于a是列向

量，那么 $a^T a$ 是一个常数，那么得到的是： $\frac{aa^T}{a^T a}$ 。所以我们可以算出投影矩阵是 $\frac{1}{9} \begin{bmatrix} 4 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 4 \end{bmatrix}$ 。

明显这是一个奇异矩阵，那么他的行列式是0。由于他的秩是1（学这么久线性代数，这个应该是可以一眼看出来的，如果不能，那么你仍然需要花一些时间去复习第一单元），那么零空间的维度是2。由于这是一个 $3 \times 3$ 矩阵，那么矩阵是会有3个特征向量的（不一定独立）。而我们之前有一个结论，矩阵的零空间的维度等于矩阵有多少个以0特征值对应的特征向量，那么就有两个特征值为0。还剩下一个特征值，可知特征值的和等于矩阵的迹，那么还剩下一个特征值是1。然后我们代入矩阵中求得特征向量，由于前面两个特征向量是在零空间中，所以本质上是求零空间的基，那么我们就不要去带大家计算了。这里讨论最后一个特征向量：首先当矩阵的特征值是1时，这个矩阵还刚好是投影向量！可以知道的是当这个矩阵对 $a$ 投影时，结果就是他本身，那么这个 $a$ 就符合特征向量！所以最后一个特征向量是 $a$ ！

顺着这个题目进行扩展，对差分方程进行求解，当我们有方程： $U_{k+1} = PU_k$ ， $U_0 = \begin{bmatrix} 9 \\ 9 \\ 0 \end{bmatrix}$ 那么求

出 $U_k$ ？

这个题目我们应该也是轻车熟路，然而一切都因为 $P$ 是投影矩阵而变得妙不可言！

当我们求解 $U_k$ 时，我们第一反应就是求解 $P$ 的特征向量与特征值，然后套用通解形式

$U_k = C_1 \lambda_1^k x_1 + C_2 \lambda_2^k x_2 + C_3 \lambda_3^k x_3$ 。或者用 $U_k = S \Lambda^k S^{-1} U_0$ ，但是大家是否记得我们是怎么推导出前面两个的公式的吗，都是基于 $U_k = P^k U_0$ 。而由于 $P$ 是投影矩阵，那么矩阵的 $N$ 次方都是等于矩阵本身的（投影矩阵的性质！），那么在这个题目中： $U_k = P U_0$ ，那么这么一看好像就是要

把 $U_0$ 投影到 $a$ 上，那么 $U_k = \frac{aa^T}{a^T a} U_0 = \begin{bmatrix} 6 \\ 3 \\ 6 \end{bmatrix}$ 。

## (2)

给定一组点（1,4）（2，5）（3,8），尝试将其拟合到一条过原点的直线上。这个问题我们在2.3也是练过好多次的。

解：

先设直线 $y = Dt$ ，只有一个未知数，列出矩阵形式：

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} D = \begin{bmatrix} 4 \\ 5 \\ 8 \end{bmatrix}$$

然后求最优解的公式是： $A^T A \hat{x} = A^T b$ 。而 $A$ 就等于 $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ ，求解得到 $D = \frac{38}{14}$ 。

我们还是要强调最小二乘法求最优解的东西是，一方面是希望大家可以有一个拟合的图形在脑海中，这在机器学习，统计学中都是常用的。第二方面是大家要记住求最优解就是一个投影矩阵把要操作的 $b$ 投影到 $A$ 上，投影的结果就在 $A$ 的列空间中，这会是最靠近 $b$ 的解！希望大家有这两方面的了解！

### (3)

有两个向量： $a_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ ,  $a_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ , 求解一组正交基？

直接使用Gram-Schmidt 正交化方法，这时我们以 $a_1$ 为模板得到的是：

$$B = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{a_1^T a_2}{a_1^T a_1} a_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{6}{14} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

### (4)

给定一个 4x4 矩阵，其特征值为 $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ 。

(a) 什么情况下矩阵是可逆矩阵。

如果矩阵是可逆的，那么矩阵一定是满秩的，也就是说矩阵的零空间只有零向量，那么特征值就不能出现0！

(b) 什么情况下矩阵的特征矩阵维度是4。

我们之前推导过，有多少个不同值的特征值，就有多少个独立的特征向量！所以特征值得各不相同！

(c)  $|A^{-1}|$ 的行列式等于多少？

还记得我们在2.8讨论课说到过的一个二级结论吗，设 $p(t)$ 是一个多项式，若 $x$ 是矩阵 $A$ 对应特征值 $\lambda$ 的特征向量，则有  $p(A)x = p(\lambda)x$ ，所以当矩阵进行倒数时，对应的特征值也要倒数。而行列式等于特征值的积，那么特征值为  $\frac{1}{\lambda_1 \lambda_2 \lambda_3 \lambda_4}$ 。

(d) 矩阵  $A + I$  的迹是多少？

我们知道迹等于矩阵特征值之和。同样运用行列式的二级结论，那么矩阵 $A+I$ 的特征值等于 $A$ 的特征值+1。然后呢就得到迹是 $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + 4$ 。

### (5)

有一个 $A_4$ 矩阵为： $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$ ，那么矩阵 $A_n$ 的行列式 $D_n$ 是多少？

解：矩阵 $A_n$ 有 $n$ 行 $n$ 列。要求他的行列式，如果要硬求，几乎是没有什么希望的。但是我们可以推导出一个递推公式（因为矩阵之间有规律！）

那么我们根据代数余子式的知识，，类比于此类矩阵在四阶时的特征，不难得到：

$$D_n = D_{n-1} - D_{n-2}$$

然后我们就可以把这个二阶递推公式化为一阶的：

$$\begin{bmatrix} D_n \\ D_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} D_{n-1} \\ D_{n-2} \end{bmatrix}$$

这个构造过程在前几讲都有推导过多次，这里不重复！

这个公式出来，我们就可以通过找矩阵  $\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$  特征值特征向量来找  $D_n$ 。

回顾我们解决问题的过程，绝大部分都是在上课讲过的内容，但是关键部分转化为递推公式的部分比较难以想到！

而还有值得一提的就是特征值为复数的情况下是怎么样：

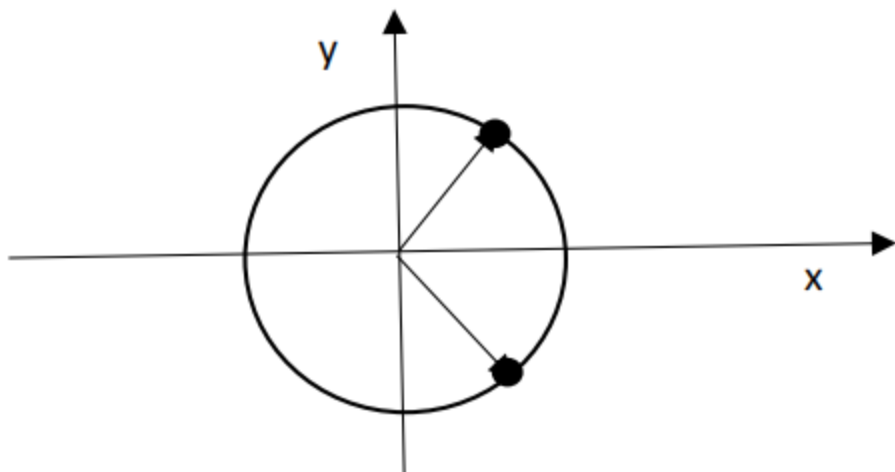
看我们矩阵  $\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$ ，那么求出矩阵的特征值得到：

$$\lambda = \frac{1 \pm \sqrt{3}i}{2}$$

特征值是两个复数，在之前我们讨论过不同特征值情况下的稳态问题，而在这我们要深入讨论一下当特征值为复数的时候的情况。

当我们画一个直角坐标系作为复平面！我们有两个点： $(\frac{1}{2}, \frac{\sqrt{3}}{2})$ ,  $(\frac{1}{2}, -\frac{\sqrt{3}}{2})$ 。观察可得，这两个点的模长都是1，而且都是单位根（即存在某个正整数  $k$ ，使得  $\lambda^k = 1$ ）。

画出图像是：



他们对应的角度都是60度。而特征值每乘一次本身，就转动60度，那么刚好6次一个周期，计算6次刚好得到1（大家可以验算看看！）。那么当特征值是原来的六次方了，显然矩阵也变成了原来的6次方。那么矩阵的6次方就是单位矩阵（首先特征值的和是2，特征值的积是，根据之前的二级结论那么矩阵只能是单位矩阵！）。

所以这个矩阵是一个周期震荡矩阵，即不收敛也不发散！

把“三角形式”换成“指数形式”： $e^{i\theta} = \cos\theta + i\sin\theta$ 。这个公式就是大名鼎鼎的欧拉公式！

tips：注意，不是所有特征值为复数的矩阵都是周期性的！要满足两个条件：

1. 所有特征值的模都等于 1（位于单位圆上）。
2. 这些特征值都是单位根（即存在某个正整数  $k$ ，使得  $\lambda^k = 1$ ）。

两个条件的关系是：

情形	例子	行为
模 = 1，但不是单位根	$\lambda = e^{i\sqrt{2}}$	<b>准周期</b> （近似周期，永不严格重复）
模 < 1	$\lambda = 0.9 e^{i\theta}$	<b>螺旋收敛到 0</b>
模 > 1	$\lambda = 1.1 e^{i\theta}$	<b>螺旋发散到 <math>\infty</math></b>

周期性矩阵在现代工程控制理论和动力系统稳定性分析中有重要作用。

## (6)

有一种矩阵：

$$A_4 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 0 & 2 & 0 & 3 \\ 0 & 0 & 3 & 0 \end{bmatrix}$$

类似的  $A_3 = A_3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 0 \end{bmatrix}$ 。

(a)：求投影到  $A_3$  列空间的投影矩阵，把  $A_3$  矩阵代入到公式： $A(A^T A)^{-1} A^T$ 。这里要注意一点， $A_3$  是不可逆矩阵，其列空间是一个平面，所以投影矩阵实际上是将另一个空间投影到一个平面上。下面我们给出这样矩阵的计算过程：

$$A_3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 0 \end{bmatrix}$$

可以知道的是第 3 列 = 2 × 第 1 列，秩为 2，去掉冗余列后保留前两列：

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 2 \end{bmatrix}_{3 \times 2}$$

### 2. 计算 $A^T A$

$$A^T A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}$$

### 3. 求逆

$$(A^T A)^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{5} \end{bmatrix}$$

#### 4. 计算 $A(A^T A)^{-1}$

$$A(A^T A)^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{5} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{5} \\ 1 & 0 \\ 0 & \frac{2}{5} \end{bmatrix}$$

#### 5. 右乘 $A^T$ 得投影矩阵

$$P = \begin{bmatrix} 0 & \frac{1}{5} \\ 1 & 0 \\ 0 & \frac{2}{5} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 5 & 0 \\ 2 & 0 & 4 \end{bmatrix}$$

### 最终结果

$$P = \frac{1}{5} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 5 & 0 \\ 2 & 0 & 4 \end{bmatrix}$$

(b) 求  $A_3$  的特征值和特征向量

令行列式  $|A_3 - \lambda I| = 0$ , 解得特征值为  $0, \pm\sqrt{5}$ 。

再代入  $(A - \lambda I)x = 0$  可解出特征向量。

(c)

求投影到  $A_4$  列空间的投影矩阵

这个还是代入公式求解。但是有一个简便的地方是,  $A_4$  是可逆矩阵, 是满秩的, 那么对应的投影矩阵就是单位矩阵  $I$ 。这个在对应章节推导过的!

tips:  $\triangle$  此外, 结合  $A_3$  可提出猜想: 此种规律的矩阵, 奇数号矩阵都是奇异 (不可逆) 的, 偶数号矩阵都是可逆的。

## 讨论课

这节讨论课主要是在讲述一写计算的简便方法, 建议大家去看原视频!