

一、知识概要

本节为习题课，主要回顾了 14~24 课的学习内容，并通过习题进行复习。

二. 复习

2.1 回顾知识

2.1.1. 投影部分:

首先，我们学习了正交性，给出了矩阵 $Q = [q_1 \ \dots \ q_n]$ ，当 q 都是标准正交基时，我们称 Q 是正交矩阵，此时有性质： $QQ^T = I$ 只有自己乘以自己才为1，其余均为0

然后我们还学习了投影，进而解决了 $Ax = b$ 问题，使用最小二乘拟合，当方程 $Ax = b$ 无解时，寻找最优解。

我们还介绍了 Gram-Schmidt 正交化方法，将线性无关的向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。

2.1.2. 行列式部分:

首先介绍了行列式的十个性，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。另外，行列式展开式有 $n!$ 项，同时展开时注意符号问题。

还介绍了代数余子式公式，这也是我们计算行列式的简便方法之一。这也让我们得到了逆矩阵公式： $A^{-1} = \frac{1}{|A|} C^T$ 。

2.1.3. 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。

2.2 例题

【例 1】现有 $a = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$ ，试找到向量 a 所在直线的投影矩阵并讨论其性质：

（即对于任意的向量 b ，找到可以将其投影到该直线上的投影矩阵）

解：

根据投影矩阵的公式：

$$P = A(A^T A)^{-1} A^T$$

可得投影矩阵为：

$$P = \frac{1}{9} \begin{bmatrix} 4 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 4 \end{bmatrix}$$

很明显这是一个不可逆矩阵，第一列，第三列均是第二列 $\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$ 的 2 倍，所以这个矩阵 P 秩为 1，列空间为 1 维。所以对应的零空间为 2 维的，这意味着它的两个特征值都为 0，最后一个特征值由矩阵的迹可得为： $\frac{1}{9}(4+1+4) = 1$ 。

求特征值 1 对应的特征向量：

$$Px = x$$

由于 P 是投影矩阵，所以联想实际意义，只有 P 作用于 a 向量上时，可以得到 $Pa = a$ ，此时 P 矩阵对 a 无影响。所以其对应特征向量为 a。

【延伸】

将矩阵 P 引入差分方程，题目如下：

$$\text{方程： } u_{k+1} = Pu_k, \text{ 其对应初值为 } u_0 = \begin{bmatrix} 9 \\ 9 \\ 0 \end{bmatrix}, \text{ 求解 } u_k ?$$

解：

首先看一下这个方程中的 P 是否有什么特殊性质，很明显，P 代表投影，这样的话 P^k 与 P 单独作用在一个向量上的效果并没有不同。所以，现在求出投影一次后的 u_0 至关重要：

$$\text{将 P 打开，利用已知 } a = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \text{ 求得：}$$

$$u_1 = Pu_0 = a \frac{a^T u_0}{a^T a} = a \frac{27}{9} = \begin{bmatrix} 6 \\ 3 \\ 6 \end{bmatrix}$$

由于矩阵 P 是投影矩阵，有结论：

$$u_k = P^k u_0 = P u_0 = \begin{bmatrix} 6 \\ 3 \\ 6 \end{bmatrix}$$

复习一下之前我们学习差分方程时， P 矩阵不是投影矩阵时，我们是使用特征值特征向量展开来求解的：

$$u_0 = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n$$

$$u_k = A^k u_0 = c_1 \lambda_1^k x_1 + c_2 \lambda_2^k x_2 + \cdots + c_n \lambda_n^k x_n$$

而在本题中使用这个公式：对于这里的投影矩阵 P ，有两个特征值为 0，一个特征值为 1，于是前两个 $c_1 \lambda_1^k x_1 + c_2 \lambda_2^k x_2$ 可以舍弃，最后一部分的 $c_3 x_3$ 可以通过初值 u_0 以及三个特征向量 x_1, x_2, x_3 来确定。

【例 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

解：

先设直线 $y = Dt$ ，只有一个未知数，一个自由度

列出矩阵形式：

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

记作： $AD=b$

下面求解最优的 D ：

之前学习的内容，最优方程为：

$$A^T A \hat{D} = A^T b$$

(本题中 $AD = b$ 对应前面介绍过的 $Ax = b$)

解得：

$$\hat{D} = \frac{38}{14}$$

【例 3】 已知： $a_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, a_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ ，求解一组正交基

分析：

将第一个向量设为 A，求解正交于其的一个向量 B。

以 a_2 作为模板，可以将 B 表示为 a_2 减去其在 a_1 上的投影的形式——这样就可以保证 B 与 a_1 (即 A) 相垂直了

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

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

(1) 特征值在满足什么条件下，矩阵是可逆的？

答：

根据我们之前的学习，只有特征值均不为 0 时，矩阵才可逆。否则的话 $Ax = 0x$ 就会有非零解。

行列式等于特征值相乘

(2) $|A^{-1}|$ 的行列式等于多少？

答：

根据 $|A||A^{-1}| = 1$ 以及特征值之积为 A 矩阵行列式的值这两个定理，可知

$$|A^{-1}| = \frac{1}{\lambda_1 \lambda_2 \lambda_3 \lambda_4}$$

(3) 矩阵 $A + I$ 的迹是多少？（迹等于矩阵特征值之和）

答：

I 为四阶单位阵，加上 I 使得 A 的迹在基础之上加 4。结果为：

$$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + 4$$

【例 5】已知一矩阵： $A_4 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$ ，求 A_n 的行列式 D_n ？

解：

利用代数余子式的知识，类比于此类矩阵在四阶时的特征，不难得到：

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

（如果忘了怎么推导出来的话见 19 课末例题）

19 课中我们只讨论到了上面这个递归式，随着我们的进一步学习，其实差分方程在这里也可以有所作为，下面我们将这个二阶递归式表示为方程：

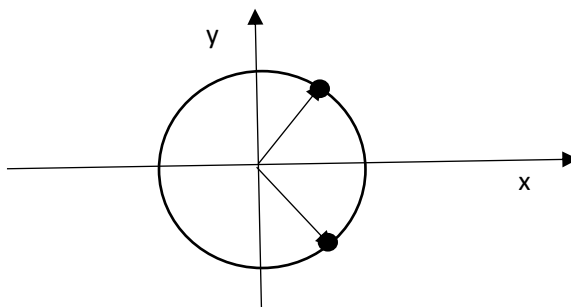
$$\begin{bmatrix} D_n \\ D_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} D_{n-1} \\ D_{n-2} \end{bmatrix}$$

（这个构造过程这里不再细讲，见 22 课裴波那切数列例题构造方法）

利用 $|A - \lambda I| = 0$ ，求 $\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$ 矩阵的特征值：

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

观察这两个特征值，发现其模均为 1，在复平面上表示位于单位圆上的两个点，示意图如下：



由欧拉公式：

$$e^{i\theta} = \cos \theta + i \sin \theta$$

还可以求得对应的表示为： $e^{\frac{\pi}{3}i}$ 、 $e^{-\frac{\pi}{3}i}$ 。

【问题】矩阵本身以及特征值与稳定性有何联系？

分析：

矩阵 $A = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$ 的 6 次幂的特征值分别是 1 和 1，因此 A 的 6 次方等于单位阵 I

或者说，这些特征值的 6 次幂均等于 1。这也解释了为什么这个形式的矩阵 A 行列式的变化周期为 6。因此该数列周期变化，既不收敛，也不发散。

【例 6】有一种矩阵： $A_4 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 0 & 2 & 0 & 3 \\ 0 & 0 & 3 & 0 \end{bmatrix} = A_4^T$ ，类比： $A_3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 0 \end{bmatrix}$ 。

• (1) 求投影到 A_3 列空间的投影矩阵

解：

由投影公式：

$$P = A(A^T A)^{-1} A^T$$

代入矩阵 A_3 ，可以很快得到答案。

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。

• (2) 求 A_3 的特征值和特征向量

解：

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。

• (3) 求投影到 A_4 列空间的投影矩阵

解：

如果 A_4 矩阵可逆，则其列空间就是 R^4 。因此如果确定了投影矩阵是可逆矩阵，那么其投影矩阵很简单：即为单位阵 I，因为此时向 R^4 不影响向量本身。

A_4 矩阵对应的行列式值为 9，于是矩阵可逆。故：其投影矩阵为 4 阶单位阵

△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。

三. 学习感悟

这一章学习了很多知识，这些基本求解方法必须掌握，因为它们是我们进一步学习更深层次内容的基础。尤其是特征值的求解，投影矩阵的理解这些东西。

机器学习算法与自然语言处理公众号
丁坤博