第二十九讲：相似矩阵和若尔当形

在本讲的开始，先接着上一讲来继续说一说正定矩阵。

- 正定矩阵的逆矩阵有什么性质？我们将正定矩阵分解为$A=S\Lambda S^{-1}$，引入其逆矩阵$A^{-1}=S\Lambda^{-1}S^{-1}$，我们知道正定矩阵的特征值均为正值，所以其逆矩阵的特征值也必为正值（即原矩阵特征值的倒数）所以，正定矩阵的逆矩阵也是正定的。

- 如果$A,\ B$均为正定矩阵，那么$A+B$呢？我们可以从判定$x^T(A+B)x$入手，根据条件有$x^TAx>0,\ x^TBx>0$，将两式相加即得到$x^T(A+B)x>0$。所以正定矩阵之和也是正定矩阵。

- 再来看有$m\times n$矩阵A，则A^TA具有什么性质？我们在投影部分经常使用A^TA，这个运算会得到一个对称矩阵，这个形式的运算用数字打比方就像是一个平方，用向量打比方就像是向量的长度平方，而对于矩阵，有A^TA正定：在式子两边分别乘向量及其转置得到x^TA^TAx，分组得到$(Ax)^T(Ax)$，相当于得到了向量Ax的长度平方，则$|Ax|^2\geq 0$。要保证模不为零，则需要Ax的零空间中仅有零向量，即A的各列线性无关（$rank(A)=n$）即可保证$|Ax|^2>0$，A^TA正定。

- 另外，在矩阵数值计算中，正定矩阵消元不需要进行“行交换”操作，也不必担心主元过小或为零，正定矩阵具有良好的计算性质。

接下来进入本讲的正题。

相似矩阵

先列出定义：矩阵$A,\ B$对于某矩阵M满足$B=M^{-1}AM$时，成$A,\ B$互为相似矩阵。

对于在对角化一讲（第二十二讲）中学过的式子$S^{-1}AS=\Lambda$，则有A相似于Λ。

- 举个例子，$A=\begin{bmatrix}2&1\\1&2\end{bmatrix}$，容易通过其特征值得到相应的对角矩阵$\Lambda=\begin{bmatrix}3&0\\0&1\end{bmatrix}$，取$M=\begin{bmatrix}1&4\\0&1\end{bmatrix}$，则$B=M^{-1}AM=\begin{bmatrix}1&-4\\0&1\end{bmatrix}\begin{bmatrix}2&1\\1&2\end{bmatrix}\begin{bmatrix}1&4\\0&1\end{bmatrix}=\begin{bmatrix}-2&-15\\1&6\end{bmatrix}$。

我们来计算这几个矩阵的的特征值（利用迹与行列式的性质），$\lambda_\Lambda=3,\ 1$、$\lambda_A=3,\ 1$、$\lambda_B=3,\ 1$。

所以，相似矩阵有相同的特征值。

* 继续上面的例子，特征值为3, 1的这一族矩阵都是相似矩阵，如$\begin{bmatrix}3&7\\0&1\end{bmatrix}$、$\begin{bmatrix}1&7\\0&3\end{bmatrix}$，其中最特殊的就是$\Lambda$。

现在我们来证明这个性质，有$Ax=\lambda x,\ B=M^{-1}AM$，第一个式子化为$AMM^{-1}x=\lambda x$，接着两边同时左乘$M^{-1}$得$M^{-1}AMM^{-1}x=\lambda M^{-1}x$，进行适当的分组得$\left(M^{-1}AM\right)M^{-1}x=\lambda M^{-1}x$即$BM^{-1}x=\lambda M^{-1}x$。

$BM^{-1}=\lambda M^{-1}x$可以解读成矩阵B与向量$M^{-1}x$之积等于λ与向量$M^{-1}x$之积，也就是B的仍为λ，而特征向量变为$M^{-1}x$。

以上就是我们得到的一族特征值为3, 1的矩阵，它们具有相同的特征值。接下来看特征值重复时的情形。

* 特征值重复可能会导致特征向量短缺，来看一个例子，设$\lambda_1=\lambda_2=4$，写出具有这种特征值的矩阵中的两个$\begin{bmatrix}4&0\\0&4\end{bmatrix}$，$\begin{bmatrix}4&1\\0&4\end{bmatrix}$。其实，具有这种特征值的矩阵可以分为两族，第一族仅有一个矩阵$\begin{bmatrix}4&0\\0&4\end{bmatrix}$，它只与自己相似（因为$M^{-1}\begin{bmatrix}4&0\\0&4\end{bmatrix}M=4M^{-1}IM=4I=\begin{bmatrix}4&0\\0&4\end{bmatrix}$，所以无论$M$如何取值该对角矩阵都只与自己相似）；另一族就是剩下的诸如$\begin{bmatrix}4&1\\0&4\end{bmatrix}$的矩阵，它们都是相似的。在这个“大家族”中，$\begin{bmatrix}4&1\\0&4\end{bmatrix}$是“最好”的一个矩阵，称为若尔当形。

若尔当形在过去是线性代数的核心知识，但现在不是了（现在是下一讲的奇异值分解），因为它并不容易计算。

* 继续上面的例子，我们在在出几个这一族的矩阵$\begin{bmatrix}4&1\\0&4\end{bmatrix}$，$\begin{bmatrix}5&1\\-1&3\end{bmatrix}$，$\begin{bmatrix}4&0\\17&4\end{bmatrix}$，我们总是可以构造出一个满足$trace(A)=8,\ \det A=16$的矩阵，这个矩阵总是在这一个“家族”中。

若尔当形

再来看一个更加“糟糕”的矩阵：

- 矩阵$\begin{bmatrix}0&1&0&0\\0&0&1&0\\0&0&0&0\\0&0&0&0\end{bmatrix}$，其特征值为四个零。很明显矩阵的秩为$2$，所以其零空间的维数为$4-2=2$，即该矩阵有两个特征向量。可以发现该矩阵在主对角线的上方有两个1，在对角线上每增加一个1，特征向量个个数就减少一个。
- 令一个例子，$\begin{bmatrix}0&1&0&0\\0&0&0&0\\0&0&0&1\\0&0&0&0\end{bmatrix}$，从特征向量的数目看来这两个矩阵是相似的，其实不然。

  若尔当认为第一个矩阵是由一个3×3的块与一个1×1的块组成的$\left[\begin{array}{ccc|c}0&1&0&0\\0&0&0&0\\0&0&0&1\\\hline 0&0&0&0\end{array}\right]$，而第二个矩阵是由两个$2\times 2$矩阵组成的$\left[\begin{array}{cc|cc}0&1&0&0\\0&0&0&0\\\hline 0&0&0&1\\0&0&0&0\end{array}\right]$，这些分块被称为若尔当块。

若尔当块的定义型为$J_i=\begin{bmatrix}\lambda_i&1&&\cdots&\\&\lambda_i&1&\cdots&\\&&\lambda_i&\cdots&\\\vdots&\vdots&\vdots&\ddots&\\&&&&\lambda_i\end{bmatrix}$，它的对角线上只为同一个数，仅有一个特征向量。

所有有，每一个矩阵A都相似于一个若尔当矩阵，型为$J=\left[\begin{array}{c|c|c|c}J_1&&&\\\hline &J_2&&\\\hline &&\ddots&\\\hline &&&J_d\end{array}\right]$。注意，对角线上方还有$1$。若尔当块的个数即为矩阵特征值的个数。

在矩阵为“好矩阵”的情况下，n阶矩阵将有n个不同的特征值，那么它可以对角化，所以它的若尔当矩阵就是Λ，共n个特征向量，有n个若尔当块。