第二十一讲：特征值和特征向量

特征值、特征向量的由来

给定矩阵A，矩阵A乘以向量x，就像是使用矩阵A作用在向量x上，最后得到新的向量Ax。在这里，矩阵A就像是一个函数，接受一个向量x作为输入，给出向量Ax作为输出。

在这一过程中，我们对一些特殊的向量很感兴趣，他们在输入(x)输出(Ax)的过程中始终保持同一个方向，这是比较特殊的，因为在大多情况下，Ax与x指向不同的方向。在这种特殊的情况下，Ax平行于x，我们把满足这个条件的x成为特征向量（Eigen vector）。这个平行条件用方程表示就是：

$$Ax = \lambda x \tag{1}$$

- 对这个式子，我们试着计算特征值为0的特征向量，此时有$Ax = 0$，也就是特征值为0的特征向量应该位于A的零空间中。

  也就是说，如果矩阵是奇异的，那么它将有一个特征值为$\lambda = 0$。

- 我们再来看投影矩阵$P = A(A^TA)^{-1}A^T$的特征值和特征向量。用向量b乘以投影矩阵P得到投影向量Pb，在这个过程中，只有当b已经处于投影平面（即A的列空间）中时，Pb与b才是同向的，此时b投影前后不变（$Pb = 1 \cdot b$）。

  即在投影平面中的所有向量都是投影矩阵的特征向量，而他们的特征值均为1。

  再来观察投影平面的法向量，也就是投影一讲中的e向量。我们知道对于投影，因为$e \perp C(A)$，所以$Pe = 0e$，即特征向量e的特征值为0。

  于是，投影矩阵的特征值为$\lambda = 1, 0$。

- 再多讲一个例子，二阶置换矩阵$A = \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$，经过这个矩阵处理的向量，其元素会互相交换。

那么特征值为1的特征向量（即经过矩阵交换元素前后仍然不变）应该型为$\begin{bmatrix}1\\1\end{bmatrix}$。

特征值为-1的特征向量（即经过矩阵交换元素前后方向相反）应该型为$\begin{bmatrix}1\\-1\end{bmatrix}$。

再提前透露一个特征值的性质：对于一个$n\times n$的矩阵，将会有n个特征值，而这些特征值的和与该矩阵对角线元素的和相同，因此我们把矩阵对角线元素称为矩阵的迹（trace）。$$\sum_{i=1}^n \lambda_i=\sum_{i=1}^n a_{ii}$$

在上面二阶转置矩阵的例子中，如果我们求得了一个特征值1，那么利用迹的性质，我们就可以直接推出另一个特征值是-1。

求解$Ax=\lambda x$

对于方程$Ax=\lambda x$，有两个未知数，我们需要利用一些技巧从这一个方程中一次解出两个未知数，先移项得$(A-\lambda I)x=0$。

观察$(A-\lambda I)x=0$，右边的矩阵相当于将A矩阵平移了λ个单位，而如果方程有解，则这个平移后的矩阵$(A-\lambda I)$一定是奇异矩阵。根据前面学到的行列式的性质，则有$$\det{(A-\lambda{I})}=0\tag{2}$$

这样一来，方程中就没有x了，这个方程也叫作特征方程（characteristic equation）。有了特征值，代回$(A-\lambda I)x=0$，继续求$(A-\lambda I)$的零空间即可。

- 现在计算一个简单的例子，$A=\begin{bmatrix}3&1\\1&3\end{bmatrix}$，再来说一点题外话，这是一个对称矩阵，我们将得到实特征值，前面还有置换矩阵、投影矩阵，矩阵越特殊，则我们得到的特征值与特征向量也越特殊。看置换矩阵中的特征值，两个实数$1,-1$，而且它们的特征向量是正交的。

  回到例题，计算$\det(A-\lambda I)=\begin{vmatrix}3-\lambda&1\\1&3-\lambda\end{vmatrix}$，也就是对角矩阵平移再取行列式。原式继续化简得$(3-\lambda)^2-1=\lambda^2-6\lambda+8=0,\lambda_1=4,\lambda_2=2$。可以看到一次项系数$-6$与矩阵的迹有关，常数项与矩阵的行列式有关。

  继续计算特征向量，$A-4I=\begin{bmatrix}-1&1\\1&-1\end{bmatrix}$，显然矩阵是奇异的（如果是非奇异说明特征值计算有误），解出矩阵的零空间$x_1=\begin{bmatrix}1\\1\end{bmatrix}$；同理计算另一个特征向量，$A-$

$2I=\begin{bmatrix}1&1\\1&1\end{bmatrix}$，解出矩阵的零空间$x_2=\begin{bmatrix}1\\-1\end{bmatrix}$。

回顾前面转置矩阵的例子，对矩阵$A'=\begin{bmatrix}0&1\\1&0\end{bmatrix}$有$\lambda_1=1,x_1=\begin{bmatrix}1\\1\end{bmatrix},\lambda_2=-1,x_2=\begin{bmatrix}-1\\1\end{bmatrix}$。看转置矩阵$A'$与本例中的对称矩阵$A$有什么联系。

易得$A=A'+3I$，两个矩阵特征值相同，而其特征值刚好相差3。也就是如果给一个矩阵加上$3I$，则它的特征值会加3，而特征向量不变。这也很容易证明，如果$Ax=\lambda x$，则$(A+3I)x=\lambda x+3x=(\lambda+3)x$，所以$x$还是原来的$x$，而$\lambda$变为$\lambda+3$。

接下来，看一个关于特征向量认识的误区：已知$Ax=\lambda x,Bx=\alpha x$，则有$(A+B)x=(\lambda+\alpha)x$，当$B=3I$时，在上例中我们看到，确实成立，但是如果B为任意矩阵，则推论不成立，因为这两个式子中的特征向量x并不一定相同，所以两个式子的通常情况是$Ax=\lambda x,By=\alpha y$，它们也就无从相加了。

- 再来看旋转矩阵的例子，旋转90°的矩阵$Q=\begin{bmatrix}\cos 90&-\sin 90\\\sin 90&\cos 90\end{bmatrix}=\begin{bmatrix}0&-1\\1&0\end{bmatrix}$（将每个向量旋转$90^\circ$，用$Q$表示因为旋转矩阵是正交矩阵中很重要的例子）。

  上面提到特征值的一个性质：特征值之和等于矩阵的迹；现在有另一个性质：特征值之积等于矩阵的行列式。$$\prod_{i=1}^n\lambda_i=\det A$$

  对于Q矩阵，有$\begin{cases}\lambda_1+\lambda_2&=0\\\lambda_1\cdot\lambda_2&=1\end{cases}$，再来思考特征值与特征向量的由来，哪些向量旋转$90^\circ$后与自己平行，于是遇到了麻烦，并没有这种向量，也没有这样的特征值来满足前面的方程组。

  我们来按部就班的计算，$\det(Q-\lambda I)=\begin{vmatrix}\lambda&-1\\1&\lambda\end{vmatrix}=\lambda^2+1=0$，于是特征值为$\lambda_1=i,\lambda_2=-i$，我们看到这两个值满足迹与行列式的方程组，即使矩阵全是实数，其特征值也可能不是实数。本例中即出现了一对共轭负数，我们可以说，如果矩阵越接近对称，那么特征值就是实数。如果矩阵越不对称，就像本例，$Q^T=-Q$，这是一个反对称的矩阵，于是我得到了纯虚的特征值，这是极端情况，通常我们见到的矩阵是介于对称与反对称之间的。

  于是我们看到，对于好的矩阵（置换矩阵）有实特征值及正交的特征向量，对于不好的矩阵（90°旋转矩阵）有纯虚的特征值。

- 再来看一个更糟的情况，$A=\begin{bmatrix}3&1\\0&3\end{bmatrix}$，这是一个三角矩阵，我们可以直接得出其特征值，即对角线元素。来看如何得到这一结论的：$\det(A-\lambda I)=\begin{vmatrix}3-\lambda&1\\0&3-\lambda\end{vmatrix}=(3-\lambda)^2=0$，于是$\lambda_1=3,\lambda_2=3$。而我们说这是一个糟糕的状况，在于它的特征向量。

  带入特征值计算特征向量，带入$\lambda_1=3$得$(A-\lambda I)x=\begin{bmatrix}0&1\\0&0\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$，算出一个特征值$x_1=\begin{bmatrix}1\\0\end{bmatrix}$，当我们带入第二个特征值$\lambda_1=3$时，我们无法得到另一个与$x_1$线性无关的特征向量了。

  而本例中的矩阵A是一个退化矩阵（degenerate matrix），重复的特征值在特殊情况下可能导致特征向量的短缺。

这一讲我们看到了足够多的“不好”的矩阵，下一讲会介绍一般情况下的特征值与特征向量。