第二十四讲：马尔科夫矩阵、傅里叶级数

马尔科夫矩阵

马尔科夫矩阵（Markov matrix）是指具有以下两个特性的矩阵：

1. 矩阵中的所有元素大于等于0；（因为马尔科夫矩阵与概率有关，而概率是非负的。）
2. 每一列的元素之和为1

对于马尔科夫矩阵，我们关心幂运算过程中的稳态（steady state）。与上一讲不同，指数矩阵关系特征值是否为0，而幂运算要达到稳态需要特征值为1。

根据上面两条性质，我们可以得出两个推论：

1. 马尔科夫矩阵必有特征值为1；
2. 其他的特征值的绝对值皆小于1。

使用第二十二讲中得到的公式进行幂运算$u_k=A^ku_0=S\Lambda^kS^{-1}u_0=S\Lambda^kS^{-1}Sc=S\Lambda^kc=c_1\lambda_1^kx_1+c_2\lambda_2^kx_2+\cdots+c_n\lambda_n^kx_n$，从这个公式很容易看出幂运算的稳态。比如我们取$\lambda_1=1$，其他的特征值绝对值均小于$1$，于是在经过$k$次迭代，随着时间的推移，其他项都趋近于$0$，于是在$k\to\infty$时，有稳态$u_k=c_1x_1$，这也就是初始条件$u_0$的第$1$个分量。

我们来证明第一个推论，取$A=\begin{bmatrix}0.1&0.01&0.3\\0.2&0.99&0.3\\0.7&0&0.4\end{bmatrix}$，则$A-I=\begin{bmatrix}-0.9&0.01&0.3\\0.2&-0.01&0.3\\0.7&0&-0.6\end{bmatrix}$。观察$A-I$易知其列向量中元素之和均为$0$，因为马尔科夫矩阵的性质就是各列向量元素之和为$1$，现在我们从每一列中减去了$1$，所以这是很自然的结果。而如果列向量中元素和为$0$，则矩阵的任意行都可以用“零减去其他行之和”表示出来，即该矩阵的行向量线性相关。

用以前学过的子空间的知识描述，当n阶方阵各列向量元素之和皆为1时，则有$\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}$在矩阵$A-I$左零空间中，即$(A-I)^T$行向量线性相关。而$A$特征值1所对应的特征向量将在$A-I$的零空间中，因为$Ax=x\longrightarrow(A-I)x=0$。

另外，特征值具有这样一个性质：矩阵与其转置的特征值相同。因为我们在行列式一讲了解了性质10，矩阵与其转置的行列式相同，那么如果$\det(A-\lambda I)=0$，则有$\det(A-\lambda I)^T=0$，根据矩阵转置的性质有$\det(A^T-\lambda I^T)=0$，即$\det(A^T-\lambda I)=0$。这正是A^T特征值的计算式。

然后计算特征值$\lambda_1=1$所对应的特征向量，$(A-I)x_1=0$，得出$x_1=\begin{bmatrix}0.6\\33\\0.7\end{bmatrix}$，特征向量中的元素皆为正。

接下来介绍马尔科夫矩阵的应用，我们用麻省和加州这两个州的人口迁移为例：

$\begin{bmatrix}u_{cal}\\u_{mass}\end{bmatrix}_{k+1}\begin{bmatrix}0.9&0.2\\0.1&0.8\end{bmatrix}\begin{bmatrix}u_{cal}\\u_{mass}\end{bmatrix}_k$，元素非负，列和为一。这个式子表示每年有10的人口从加州迁往麻省，同时有20的人口从麻省迁往加州。注意使用马尔科夫矩阵的前提条件是随着时间的推移，矩阵始终不变。

设初始情况$\begin{bmatrix}u_{cal}\\u_{mass}\end{bmatrix}_0=\begin{bmatrix}0\\1000\end{bmatrix}$，我们先来看第一次迁徙后人口的变化情况：$\begin{bmatrix}u_{cal}\\u_{mass}\end{bmatrix}_1=\begin{bmatrix}0.9&0.2\\0.1&0.8\end{bmatrix}\begin{bmatrix}0\\1000\end{bmatrix}=\begin{bmatrix}200\\800\end{bmatrix}$，随着时间的推移，会有越来越多的麻省人迁往加州，而同时又会有部分加州人迁往麻省。

计算特征值：我们知道马尔科夫矩阵的一个特征值为$\lambda_1=1$，则另一个特征值可以直接从迹算出$\lambda_2=0.7$。

计算特征向量：带入$\lambda_1=1$求$A-I$的零空间有$\begin{bmatrix}-0.1&0.2\\0.1&-0.2\end{bmatrix}$，则$x_1=\begin{bmatrix}2\\1\end{bmatrix}$，此时我们已经可以得出无穷步后稳态下的结果了。$u_\infty=c_1\begin{bmatrix}2\\1\end{bmatrix}$且人口总数始终为1000，则$c_1=\frac{1000}{3}$，稳态时$\begin{bmatrix}u_{cal}\\u_{mass}\end{bmatrix}_\infty=\begin{bmatrix}\frac{2000}{3}\\\frac{1000}{3}\end{bmatrix}$。注意到特征值为1的特征向量元素皆为正。

为了求每一步的结果，我们必须解出所有特征向量。带入$\lambda_2 = 0.7$求$A - 0.7I$的零空间有$\begin{bmatrix} 0.2 & 0.2 \\ 0.1 & 0.1 \end{bmatrix}$，则$x_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$。

通过u_0解出c_1, c_2，$u_k = c_1 1^k \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 0.7^k \begin{bmatrix} -1 \\ 1 \end{bmatrix}$，带入$k = 0$得$u_0 = \begin{bmatrix} 0 \\ 1000 \end{bmatrix} = c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$，解出$c_1 = \frac{1000}{3}, c_2 = \frac{2000}{3}$。

另外，有时人们更喜欢用行向量，此时将要使用行向量乘以矩阵，其行向量各分量之和为1。

傅里叶级数

在介绍傅里叶级数（Fourier series）之前，先来回顾一下投影。

设$q_1, q_2, \cdots q_n$为一组标准正交基，则向量v在该标准正交基上的展开为$v = x_1 q_1 + x_2 q_2 + \cdots + x_n q_n$，此时我们想要得到各系数$x_i$的值。比如求$x_1$的值，我们自然想要消掉除$x_1 q_1$外的其他项，这时只需要等式两边同乘以$q_1^T$，因为的$q_i$向量相互正交且长度为1，则$q_i^T q_j = 0, q_i^2 = 1$所以原式变为$q_1^T v = x_1$。

写为矩阵形式有$\begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = v$，即$Qx = v$。所以有$x = Q^{-1} v$，而在第十七讲我们了解到标准正交基有$Q^T = Q^{-1}$，所以我们不需要计算逆矩阵可直接得出$x = Q^T v$。此时对于$x$的每一个分量有$x_i = q_i^T v$。

接下来介绍傅里叶级数。先写出傅里叶级数的展开式：

$$f(x) = a_0 + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \cdots$$

傅里叶发现，如同将向量v展开（投影）到向量空间的一组标准正交基中，在函数空间中，我们也可以做类似的展开。将函数$f(x)$投影在一系列相互正交的函数中。函数空间中的$f(x)$就是向量空间中的v；函数空间中的$1, \cos x, \sin x, \cos 2x, \sin 2x, \cdots$就是向量空间中的$q_1, q_2, \cdots, q_n$；不同的是，函数空间是无限维的而我们以前接触到的向量空间通常是有限维的。

再来介绍何为“函数正交”。对于向量正交我们通常使用两向量内积（点乘）为零判断。我们知道对于向量v, w的内积为$v^T w = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n = 0$，也就是向量的

每个分量之积再求和。而对于函数$f(x)\cdot g(x)$内积，同样的，我们需要计算两个函数的每个值之积而后求和，由于函数取值是连续的，所以函数内积为：

$$f^Tg=\int f(x)g(x)\mathrm{d}x$$

在本例中，由于傅里叶级数使用正余弦函数，它们的周期都可以算作2π，所以本例的函数点积可以写作$f^Tg=\int_0^{2\pi}f(x)g(x)\mathrm{d}x$。我来检验一个内积$\int_0^{2\pi}\sin x\cos x\mathrm{d}x=\frac{1}{2}\sin^2x\Big|_0^{2\pi}=0$，其余的三角函数族正交性结果可以参考傅里叶级数的“希尔伯特空间的解读”一节。

最后我们来看$\cos x$项的系数是多少（a_0是$f(x)$的平均值）。同向量空间中的情形一样，我们在等式两边同时做$\cos x$的内积，原式变为$\int_0^{2\pi}f(x)\cos x\mathrm{d}x=a_1\int_0^{2\pi}\cos^2x\mathrm{d}x$，因为正交性等式右边仅有$\cos x$项不为零。进一步化简得$a_1\pi=\int_0^{2\pi}f(x)\cos x\mathrm{d}x\rightarrow a_1=\frac{1}{\pi}\int_0^{2\pi}f(x)\cos x\mathrm{d}x$。

于是，我们把函数$f(x)$展开到了函数空间的一组标准正交基上。