线性代数

-24 课 马尔科夫矩阵，傅里叶级数

一、知识概要

本节介绍了马尔科夫矩阵与傅里叶级数，是对特征值以及之前大量知识的一次总结，重在理解应用，但是内容不是非常细致，主要是大体的了解。

二．马尔科夫矩阵的性质及其应用

2.1 马尔科夫矩阵的性质

先给出一个马尔科夫矩阵如下：

【例】

$$A=\begin{bmatrix}0.1 & 0.01 & 0.3\\ 0.2 & 0.99 & 0.3\\ 0.7 & 0 & 0.4\end{bmatrix}$$

马尔科夫矩阵要满足两条性质：

（1）每个元素均为非负数

（2）每列的元素和为 1

性质：如果 A 是马尔可夫矩阵，则 A 的幂也是马尔科夫矩阵。

接下来讨论的是矩阵次幂运算的“稳态”概念，联想上一节中学习过的微分方程中的稳态问题，那时我们是用特征值和 0 的关系来判断的，而这里的运算变为了幂次运算，这时稳态的判断就要变化了。

这就涉及到了特征值与特征向量，研究马尔科夫矩阵的特征值，发现它的特征值有这样的性质：

（1） $\lambda = 1$ 是一个特征值

（2） 其余特征值，$|\lambda_i| < 1$

其实要证明性质（1）很简单，直接代入$\lambda = 1$ 就好了，如上面的例子，得到：$A-1I=\begin{bmatrix}-0.9 & 0.01 & 0.3\\ 0.2 & -0.01 & 0.3\\ 0.7 & 0 & -0.6\end{bmatrix}$，每一列元素之和为 0，这是个不可逆矩阵，三行相加得到零行，行/列向量线性相关，所以 1 是其特征值，存在对应特征向量。或者说**每列的元素和为 1 这个性质决定了其有特征值 1 这个性质**。性质（2）课上没有证明，但是它也是成立的。

好的，回到我们的主题，判断次幂运算的收敛性。记得之前差分方程中我们接触过这样的算式：$u_k = A^k u_0$，这也可以看做矩阵次幂运算的问题。

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

根据马尔科夫矩阵的特点，$\lambda_1 = 1$，其余$|\lambda_i| < 1$，所以最后次幂运算会收敛于u_0中的c_1x_1部分，这就是其稳态性质：马尔科夫矩阵中，若其特征值 1 所对应的特征向量的元素全为正，则初始值u_0是正的，那么稳态也是正值。

这里还引入了一个对之前转置矩阵特征值的补充：
由之前介绍过的行列式性质十，转置矩阵与原矩阵行列式值相同，具体这里应用到 A-λi 这个方程上，得到：

$$|A-\lambda i| = |(A-\lambda i)^T| = |A^T-\lambda i|$$

也就是说，$|A-\lambda i| = 0$ 时，$|A^T-\lambda i|$也为 0，它们的特征值是一样的。但是另外还要注意，虽然这样能得到：A 的特征值 $= A^T$的特征值，但是它们的特征向量并不一样，因为根本不属于 A 的同一个空间中。

2.2 马尔科夫矩阵的应用

下面研究方程：$u_{k+1} = Au_k$（其中，A 是马尔科夫矩阵）

给定 A 是 2x2 的矩阵，研究加州和麻省的人口问题，矩阵 A 表示，一年后发生了人口的迁移。矩阵的四个元素相应表示的是留下和迁出的概率，在这个过程中总人口数保持不变。这里有个严格的限制：那就是马尔科夫矩阵不变（即每次变动的概率一样）。
给定等式如下形式：

$$\begin{bmatrix} U_{加} \\ U_{麻} \end{bmatrix}_{t=k+1} = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} U_{加} \\ U_{麻} \end{bmatrix}_{t=k}$$

（中间的矩阵表示有 0.9 的比例的人口留在加州，0.1 从加州迁移到麻省，0.8 的人留在麻省，0.2 的人迁移到加州；这是一个马尔科夫矩阵）

下面分析其稳态：
现在假定初始状态为：

$$\begin{bmatrix} U_{加} \\ U_{麻} \end{bmatrix}_{t=0} = \begin{bmatrix} 0 \\ 1000 \end{bmatrix}$$

经过一次变迁之后，两州的人口状况是：

$$\begin{bmatrix} U_{加} \\ U_{麻} \end{bmatrix}_{t=1} = \begin{bmatrix} 200 \\ 800 \end{bmatrix}$$

计算该马尔科夫矩阵特征值为 1 和 0.7。对应求得特征向量为$\begin{bmatrix}2\\1\end{bmatrix}$，$\begin{bmatrix}-1\\1\end{bmatrix}$

现在我们来看经过无穷次迁移之后，人口状况的稳态是怎样的？
根据公式：

$$u_k = c_1 1^k \begin{bmatrix}2\\1\end{bmatrix} + c_2 (0.7)^k \begin{bmatrix}-1\\1\end{bmatrix}$$

只有$\begin{bmatrix}2\\1\end{bmatrix}$会反映最后稳态时的人口分布状况。

代入初始情况计算 c 的值：

$$u_0 = \frac{1000}{3}\begin{bmatrix}2\\1\end{bmatrix} + \frac{2000}{3}\begin{bmatrix}-1\\1\end{bmatrix} = \begin{bmatrix}0\\1000\end{bmatrix}$$

最后得到的就是：$\frac{1000}{3}\begin{bmatrix}2\\1\end{bmatrix}$。总数不变。分布如该$u_0$中的$c_1 x_1$反映出来。

机器学习算法与自然语言处理公众号
丁坤博

三．傅里叶级数

3.1 前提基础

假定有一组(n 个)标准正交向量$q_1, q_2, \ldots q_n$，它们是 n 维空间一组完整基。那么此空间中，任意向量 v 可由这组基的线性组合表示：

$$v = x_1 q_1 + x_2 q_2 + \ldots + x_n q_n$$

矩阵化的表示方法是：

$$Qx = v$$

于是有：

$$x = Q^{-1} v$$

由于 Q 是正交阵，之前介绍过有$Q^T = Q^{-1}$

可以得到：

$$x = Q^{-1} v = Q^T v$$

对应到各个分量：

$$q_i^T v = x_i$$

这个结论还可以用之前的$v = x_1 q_1 + x_2 q_2 + \ldots + x_n q_n$计算：

由于$v = x_1 q_1 + x_2 q_2 + \ldots + x_n q_n$中的$q_i$都是标准正交基，由$q_i^T q_j \begin{cases} 0 & (i \neq j) \\ 1 & (i = j) \end{cases}$，在$v = x_1 q_1 + x_2 q_2 + \ldots + x_n q_n$两侧同时乘上$q_i^T$，也能得到$q_i^T v = x_i$这个结果。其意义即为向量与标准正交基的点乘，求得各分量。

3.2 傅里叶级数

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

这个问题与之前的问题有什么关系？一样都是无穷维，但关键性质是**函数正交**。

注：

上述形式就是傅里叶级数，作用在函数空间上，用函数$f(x)$来代替向量 v，也就是使用正交函数来代替原本上面介绍的标准正交基 $q_1, q_1, q_3...$。反映到上面的函数中，基就是 1，cosx，sinx，cos2x，sin2x...

机器学习算法与自然语言处理公众号
丁坤博

而且傅里叶级数成立的原因即是：这些基是正交的。

(1)下面解释下**函数空间下的“正交”**：

我们知道，对于向量的点积为：

$$v^T w = v_1 w_1 + \ldots + v_n w_n$$

在这里，向量可以看做是离散的，但是上面给出的函数，它们在其定义域上是连续的，那么对于两个连续函数而言，其内积是什么？

对比上面的向量内积的形式，对于函数而言，与之最相似的情况就是在每个 x 值上的$f(x) \cdot g(x)$，而连续情况对应的就是对其进行积分。至于其积分上下限：观察到函数$f(x)$是一个周期函数，周期为2π。所以这里最好也是从 0 积分到2π。

定义好函数空间的内积之后，我们可以检验一下正交性，明显可以得到傅里叶级数中的基是正交的。

（2）下面解决系数的求解问题。

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

以求解其中的中的a_1为例，与之前讲过的向量求解是类似的，利用正交基，直接将向量与正交基点乘来求各分量。

这里将函数 f（x）与 cosx 作内积，所以就可以得到：

$$\int_0^{2\pi} f(x) cosx dx = \int_0^{2\pi} (a_0 + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x) cosx dx$$

投影到 cos 这个基上，经过计算剩下的也只剩下了$\int_0^{2\pi} a_1 cosx^2 dx = a_1$π，和向量形式一样。所以，$a_1$的值为：

$$a_1 = \frac{\int_0^{2\pi} f(x) \cos x dx}{\pi}$$

机器学习算法与自然语言处理公众号 丁坤博

其他的对应值以此类推，可见，函数也可以展开到一组标准正交基上。

三、学习感悟

这部分内容重在理解，马尔科夫矩阵的重点在于理解那个人口迁移模型的应用，而傅里叶级数实际上就是将我们之前向标准正交机投影的向量改为了空间中向函数投影的展开形式，只要理解了正交函数的概念，这些问题便是显而易见的了。