

MIT 18.06 线性代数

第二十四讲：马尔可夫链、傅里叶级数

整理自 MIT 18.06 课程及相关公开笔记

引言

本讲是特征值与投影矩阵理论在具体应用中的展现，主要内容涵盖两部分：一是马尔可夫矩阵 (Markov Matrices) 及其在人口迁移与稳态分布中的应用；二是傅里叶级数 (Fourier Series)，其核心思想是将标准正交基从有限维向量空间推广至无限维函数空间，是一种广义的正交投影。

1 马尔可夫矩阵及其性质

1.1 定义

如果一个方阵 A 满足以下两个条件，则称其为**马尔可夫矩阵** (Markov Matrix)：

1. 矩阵中的所有元素均为非负数 ($a_{ij} \geq 0$)，
2. 矩阵**每一列**的元素之和均为 1。

这种构造源于概率论的应用，矩阵中的元素常被解释为状态转移的概率。马尔可夫矩阵的一个直观性质是：对马尔可夫矩阵求任意正整数次幂 (A^k)，其结果仍然是马尔可夫矩阵。

1.2 特征值性质与稳态

马尔可夫矩阵具备两个关键的特征值属性：

1. $\lambda_1 = 1$ 必定是它的一个特征值，
2. 其余的所有特征值的绝对值均满足 $|\lambda_i| < 1$ 。

证明 $\lambda = 1$ 是特征值：马尔可夫矩阵 A 的各列元素之和为 1。考察矩阵 $(A - I)$ ，对于该矩阵的每一列，其元素之和变为 $1 - 1 = 0$ ，由于列元素之和为 0，这意味着将矩阵的所有行相加后，得到的是一个全零行，因此矩阵 $(A - I)$ 不可逆，故存在非零向量 $\mathbf{x}$ 使得 $(A - I)\mathbf{x} = \mathbf{0}$ ，即 $A\mathbf{x} = \mathbf{x}$ ， $\lambda = 1$ 是其特征值，且特征值 1。

稳态的意义：根据差分方程解的公式 $\mathbf{u}_k = A^k \mathbf{u}_0 = c_1 \lambda_1^k \mathbf{x}_1 + c_2 \lambda_2^k \mathbf{x}_2 + \cdots + c_n \lambda_n^k \mathbf{x}_n$ ，由于马尔可夫矩阵 $|\lambda_i| < 1$ ($i \geq 2$)，当 $k \rightarrow \infty$ 时，所有包含 λ_i^k 的项都将趋近于零。因此，系统最终会收敛到稳态 (steady state)：

$$\mathbf{u}_{\text{steady}} = \lim_{k \rightarrow \infty} \mathbf{u}_k = c_1 \mathbf{x}_1$$

即系统最终将稳定在初始向量 $\mathbf{u}_0$ 在特征向量 $\mathbf{x}_1$ 方向上的分量。如果 $\mathbf{x}_1$ 的分量均为正，且初始状态为正，则稳态分布也必然为正。

1.3 应用实例：加州与麻省的人口迁移

考虑一个经典的马尔可夫链应用：加利福尼亚州与马萨诸塞州的人口迁移问题。假设每年这两个州的人口互相流动，比例固定。设向量 $\mathbf{u}_k = \begin{bmatrix} u_{\text{Cal}} \\ u_{\text{Mass}} \end{bmatrix}_k$ 表示第 k 年两州的人口分布，给定状态转移矩阵 A （规定列元素之和为 1，即总人数守恒）：

$$A = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}$$

矩阵的含义为：每年有 90% 的加州人留在加州，10% 迁往麻省；有 20% 的麻省人迁往加州，80% 留在麻省。设初始人口分布为 $\mathbf{u}_0 = \begin{bmatrix} 0 \\ 1000 \end{bmatrix}$ ，即全部人口一开始均在麻省。

求该矩阵的特征值：由于矩阵列和为 1，故 $\lambda_1 = 1$ 。根据矩阵的迹 $\text{tr}(A) = 0.9 + 0.8 = 1.7$ ，可得 $\lambda_2 = \text{tr}(A) - \lambda_1 = 0.7$ ，由于 $|\lambda_2| < 1$ ，系统最终将趋于稳态。求解对应的特征向量：

- 对于 $\lambda_1 = 1$ ，解 $(A - I)\mathbf{x}_1 = \mathbf{0}$ ，得特征向量 $\mathbf{x}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ ，这表示最终稳态下，加州与麻省的人口比例将稳定在 2 : 1。
- 对于 $\lambda_2 = 0.7$ ，解 $(A - 0.7I)\mathbf{x}_2 = \mathbf{0}$ ，得特征向量 $\mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ 。

根据差分方程通解 $\mathbf{u}_k = c_1(1)^k \mathbf{x}_1 + c_2(0.7)^k \mathbf{x}_2$ ，代入 $\mathbf{u}_0$ ：

$$\mathbf{u}_0 = c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1000 \end{bmatrix} \implies c_1 = \frac{1000}{3}, \quad c_2 = \frac{2000}{3}$$

因此，系统的终态（稳态）人口分布为：

$$\mathbf{u}_{\text{steady}} = \frac{1000}{3} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2000/3 \\ 1000/3 \end{bmatrix}$$

这表示随着时间的推移，大约有 2/3 的人口居住在加州，1/3 的人口居住在麻省，这是一个与初始状态无关的固有稳态。

2 傅里叶级数

2.1 从有限维向量空间到函数空间

回顾标准正交基的概念：若有一组标准正交向量 $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n$ 构成 n 维空间的一组基，则空间中的任意向量 $\mathbf{v}$ 可以唯一地表示为这组基的线性组合：

$$\mathbf{v} = x_1 \mathbf{q}_1 + x_2 \mathbf{q}_2 + \dots + x_n \mathbf{q}_n$$

求解系数 x_i ：由于 $\mathbf{q}_i^T \mathbf{q}_j = \delta_{ij}$ ，在等式两端同时左乘 $\mathbf{q}_i^T$ ，即可得到 $x_i = \mathbf{q}_i^T \mathbf{v}$ 。

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

从矩阵角度而言，这等价于求解 $Q\mathbf{x} = \mathbf{v}$ ，由于 Q 是正交矩阵， $\mathbf{x} = Q^{-1}\mathbf{v} = Q^T \mathbf{v}$ 。

傅里叶级数是将这一思想从离散的向量空间推广到了**无穷维的函数空间**。函数空间中的一组标准正交基为三角函数： $1, \cos x, \sin x, \cos 2x, \sin 2x, \cos 3x, \sin 3x, \dots$ ；任何周期为 2π 的（满足一定条件的）函数 $f(x)$ 都可以展开为傅里叶级数：

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

2.2 函数空间的内积与正交性

在函数空间中，两个函数 $f(x)$ 和 $g(x)$ ，其内积（点积）定义由离散求和变为积分。考虑到三角函数的周期为 2π ，通常定义内积为区间 $[0, 2\pi]$ 上的定积分：

$$f^T g \triangleq \int_0^{2\pi} f(x)g(x) dx$$

利用该定义可以检验傅里叶级数基的正交性。例如，验证 $\sin x$ 与 $\cos x$ 正交：

$$\int_0^{2\pi} \sin x \cos x dx = \int_0^{2\pi} \frac{1}{2} \sin 2x dx = -\frac{1}{4} \cos 2x \Big|_0^{2\pi} = 0$$

2.3 傅里叶系数求解

通过类比标准正交基下求解系数 $x_i = \mathbf{q}_i^T \mathbf{v}$ 的思路，可以利用内积求出傅里叶系数。以求系数 a_1 为例，将 $f(x)$ 与基函数 $\cos x$ 进行内积：

$$\begin{aligned} \int_0^{2\pi} f(x) \cos x dx &= \int_0^{2\pi} (a_0 + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + \dots) \cos x dx \\ &= 0 + a_1 \int_0^{2\pi} \cos^2 x dx + 0 + 0 + \dots \\ &= a_1 \pi \end{aligned}$$

因此，傅里叶系数 a_1 的计算公式为：

$$a_1 = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos x dx$$

同理，可以求出 a_0, b_1, a_2 等其他系数，这一过程说明，傅里叶级数的本质是将任意函数向由正交三角函数构成的空间中进行**投影**。

总结

本讲展示了特征值与投影理论在实际问题中的两大典型应用：

- **马尔可夫矩阵**：利用特征值为 1 和其余特征值模长小于 1 的性质，可以快速判断系统的**稳态分布**；人口迁移的实例演示了如何通过特征向量直接将稳态构成描述出来。
- **傅里叶级数**：将标准正交基的概念从离散向量空间扩展到连续函数空间，通过定义函数的内积，可以利用投影公式快速求出傅里叶系数，体现了线性代数分析无限维空间结构的强大能力。