# 随机过程核心知识点复习大纲 (详细版)

根据课程大纲整理与扩充

2025 年 7 月 2 日

# 1 第 1 讲：随机过程简介

### 随机过程的基本概念

- **定义：** 随机过程是依赖于参数 $t$ 的一族（或一个集合）的随机变量 $\{X(t), t \in T\}$。它描述了一个系统随时间（或其他参数）演变的随机行为。
- **核心要素：**
  1. **状态空间 (State Space) $S$：** 随机变量 $X(t)$ 所有可能取值的集合。
     - 离散状态：如硬币正反面 $\{0, 1\}$，排队人数 $\{0, 1, 2, \ldots\}$。
     - 连续状态：如股价 $[0, \infty)$，温度 $(-\infty, \infty)$。
  2. **参数集 (Index Set) $T$：** 参数 $t$ 的所有可能取值的集合，通常是时间。
     - 离散时间： $T = \{0, 1, 2, \ldots\}$ 或 $\{t_0, t_1, \ldots\}$，称为**随机序列**。
     - 连续时间： $T = [0, \infty)$ 或 $(-\infty, \infty)$。
- **样本函数/实现 (Sample Path)：** 对于一个固定的样本点 $\omega \in \Omega$（即一次随机试验的结果），$X(t, \omega)$ 作为 $t$ 的函数，是过程的一次具体实现。例如，某只股票从开盘到收盘的价格曲线。

# 2 第 2 讲：概率论回顾

### 概率空间与随机变量

- **概率空间 $(\Omega, \mathcal{F}, P)$：** 任何随机现象的数学模型。
- **随机变量：** 建立在概率空间上的实值函数，将随机试验的结果映射为数值。

### 数字特征与经典极限定理

- **数字特征：**
  - **期望** $E[X]$：随机变量的加权平均，描述其中心位置。
  - **方差** $D(X)$：随机变量偏离其期望的平均平方，描述其离散程度。
  - **协方差** $\mathrm{Cov}(X, Y)$：描述两个随机变量的线性相关程度。
- **极限定理：**
  - **大数定律 (LLN)**：样本均值 $\frac{1}{n}\sum X_i$ 依概率收敛于总体期望 $E[X]$。提供了用算术平均估计期望的理论依据。
  - **中心极限定理 (CLT)**：大量独立同分布随机变量之和（或均值）的分布近似于正态分布 $N(\mu, \sigma^2)$。解释了为何正态分布在自然界中如此普遍。

### 特征函数

- **定义：** $\phi_X(u) = E[e^{iuX}]$。它是随机变量的傅里叶变换。
- **核心性质：**
  1. **唯一性：** 特征函数与概率分布一一对应。
  2. **求矩：** $E[X^k] = \frac{1}{i^k}\phi_X^{(k)}(0)$。
  3. **独立可加性：** 若 $X, Y$ 独立，则 $Z = X + Y$ 的特征函数为 $\phi_Z(u) = \phi_X(u)\phi_Y(u)$。
  4. **收敛性定理 (Lévy's Continuity Theorem)：** 随机变量序列依分布收敛的充要条件是其特征函数序列逐点收敛到一个在原点连续的函数。

## 3 第 3 讲：离散时间马尔可夫链

### 基本概念与转移概率矩阵

- **马尔可夫性 (无后效性)：** 给定现在，未来与过去相互独立。

$$P\{X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0\} = P\{X_{n+1} = j | X_n = i\}$$

- **转移概率矩阵 (TPM) $P$：** 元素 $p_{ij}$ 表示从状态 $i$ 一步转移到状态 $j$ 的概率。它是一个随机矩阵（非负且行和为 1）。
- **Chapman-Kolmogorov (C-K) 方程：** $P^{(n+m)} = P^{(n)}P^{(m)}$。这表明 n 步转移矩阵是单步转移矩阵的 n 次方：$\boldsymbol{P^{(n)} = P^n}$。

### 状态分类与基本性质

- **常返态 (Recurrent)：** 从状态 $i$ 出发，以概率 1 能返回。这意味着会**无限次**返回。
- **暂态 (Transient)：** 从状态 $i$ 出发，有正概率永远不返回。这意味着只会**有限次**返回。
- **周期性：** 状态 $i$ 的周期 $d(i)$ 是所有 “$i \to i$” 的步数的最大公约数。若 $d(i) > 1$，状态表现出周期性振荡行为。
- **不可约马氏链：** 所有状态构成一个单一的互通类。这意味着从任何状态出发都可以到达任何其他状态。
- **遍历态 (Ergodic)：** 正常返且非周期的状态。这是马氏链达到稳定状态的理想条件。

### 极限行为与平稳分布

- **平稳分布 $\pi$：** 如果初始分布是 $\pi$，那么经过一步转移后，分布仍然是 $\pi$。
  - **求解方程：** $\boldsymbol{\pi = \pi P}$ 且 $\sum_j \pi_j = 1$。

– **物理解释：** $\pi_j$ 可被解释为过程在状态 $j$ 上花费时间的长期比例。

- **遍历定理 (核心定理)：** 对于一个有限状态、不可约、非周期的马氏链，存在唯一的平稳分布 $\pi$，并且极限分布等于此平稳分布。

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j, \quad \text{对所有} i, j \in S$$

# 4 第 4 讲：泊松过程

### 泊松过程的核心特征

- **描述对象**：在一段时间内，某事件发生次数的随机过程。
- **关键参数**：发生率 $\lambda > 0$。
- **三个等价定义**：

1. **定义 1 (基于增量)：**
   – **独立增量**：在不重叠时间段内发生的事件数是相互独立的。
   – **平稳增量**：在任何长度为 $t$ 的时间段内，事件发生的次数都服从参数为 $\lambda t$ 的泊松分布。

$$P\{N(t+s) - N(s) = k\} = \frac{e^{-\lambda t}(\lambda t)^k}{k!}$$

2. **定义 2 (基于到达间隔时间)：**
   – 事件的到达间隔时间 $T_1, T_2, \ldots$ 是独立同分布的，且都服从参数为 $\lambda$ 的**指数分布**。
3. **定义 3 (基于无穷小性质)：** 在一个极小的时间段 $h$ 内，
   – 发生一次事件的概率约等于 $\lambda h$。
   – 发生两次或更多次事件的概率几乎为零 ($o(h)$)。
   – 不发生事件的概率约等于 $1 - \lambda h$。

### 泊松过程的应用与推广

- **性质：**
  – **叠加性**：两个独立的泊松过程之和仍是泊松过程，速率为两者之和。
  – **筛选性**：对一个泊松过程的每次事件以概率 $p$ 进行标记，被标记的事件构成一个新的泊松过程，速率为 $\lambda p$。
- **推广：** 复合泊松过程、非齐次泊松过程。

## 5 第 5 讲：连续时间马尔可夫链

### 定义与 Q 矩阵

- **核心思想：** 过程在状态 $i$ 的停留时间服从参数为 $q_i$ 的指数分布。指数分布的“无记忆性”保证了马尔可夫性。
- **Q 矩阵 (Infinitesimal Generator)：** 也称为转移速率矩阵，描述了状态间的瞬时转移速率。

$$Q = (q_{ij})$$

  - $q_{ij}$ $(i \neq j)$ 是从状态 $i$ 到 $j$ 的**转移速率**。在极小时间 $h$ 内，从 $i$ 转移到 $j$ 的概率约为 $q_{ij}h$。
  - $q_{ii} = -q_i = -\sum_{j \neq i} q_{ij}$。$q_i$ 是从状态 $i$ 离开的**总速率**。Q 矩阵的行和恒为 0。

### 如何计算 Q 矩阵

计算 Q 矩阵通常分为两步：确定非对角线元素，然后确定对角线元素。

**第一步：确定非对角线元素 $q_{ij}$ $(i \neq j)$** 非对角线元素 $q_{ij}$ 代表从状态 $i$ 直接转移到状态 $j$ 的“速率”。你需要分析问题的物理或业务背景，找出导致这种状态变化的事件，并确定该事件的发生率。

- **识别事件：**分析系统，找出所有可能导致状态从 $i$ 变为 $j$ 的单一事件。
- **确定速率：**这些事件通常服从泊松过程或指数分布。
  - 如果状态变化是由一个速率为 $\lambda$ 的泊松过程事件触发的，那么相应的 $q_{ij} = \lambda$。
  - 如果状态 $i$ 的停留时间服从参数为 $q_i$ 的指数分布，离开后以概率 $p_{ij}$ 转移到状态 $j$，那么 $q_{ij} = q_i \cdot p_{ij}$。
- **汇总速率：**如果有多种独立的事件都可以导致从 $i$ 转移到 $j$，它们的速率分别为 $\lambda_1, \lambda_2, \ldots$，那么总的转移速率 $q_{ij} = \lambda_1 + \lambda_2 + \ldots$。
- **零速率：**如果系统无法从状态 $i$ 一步直接转移到状态 $j$，则 $q_{ij} = 0$。

**第二步：确定对角线元素 $q_{ii}$** 对角线元素 $q_{ii}$ 代表从状态 $i$ “离开”的速率。它总是负数，其绝对值等于该行所有非对角线元素之和。

$$q_{ii} = -\sum_{j \neq i} q_{ij}$$

这个定义的物理解释是，$q_i = -q_{ii}$ 是在状态 $i$ 时，**任何一种**可能的状态转移事件发生的总速率。因此，在状态 $i$ 的停留时间服从参数为 $q_i$ 的指数分布。

**示例：M/M/1 排队系统** 考虑一个 M/M/1 排队系统，状态 $n$ 表示系统中有 $n$ 个顾客。

- **状态 $n \to n+1$ (出生)：**
  - **事件：**一个新顾客到达。
  - **速率：**顾客到达服从速率为 $\lambda$ 的泊松过程。
  - **计算：** $q_{n,n+1} = \lambda$ (对所有 $n \geq 0$)。
- **状态 $n \to n-1$ (死亡)：**
  - **事件：**一个顾客完成服务并离开。
  - **速率：**服务时间服从参数为 $\mu$ 的指数分布，即服务速率为 $\mu$。这只在系统不为空 $(n \geq 1)$ 时发生。
  - **计算：** $q_{n,n-1} = \mu$ (对所有 $n \geq 1$)。
- **其他转移：**
  - 系统不能一步跳过多个状态，所以所有其他的 $q_{ij}$ (其中 $|i-j| > 1$) 都为 0。

根据以上分析，我们填写 Q 矩阵的非对角线元素：

$$Q = \begin{pmatrix} & \lambda & 0 & 0 & \dots \\ \mu & & \lambda & 0 & \dots \\ 0 & \mu & & \lambda & \dots \\ \vdots & \vdots & \ddots & & \ddots \end{pmatrix}$$

最后，根据“行和为 0”的规则，计算对角线元素：

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \dots \\ \mu & -(\lambda+\mu) & \lambda & 0 & \dots \\ 0 & \mu & -(\lambda+\mu) & \lambda & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix}$$

### 动态方程与稳态分析

- **Kolmogorov 方程：** 描述了转移概率矩阵 $P(t)$ 随时间演化的微分方程组。

$$P'(t) = P(t)Q \quad (\text{向前方程}) \quad \text{和} \quad P'(t) = QP(t) \quad (\text{向后方程})$$

  其形式解为 $\boldsymbol{P(t) = e^{Qt}}$。
- **平稳分布 $\pi$：**
  - **求解方程：** $\boldsymbol{\pi Q = 0}$ 且 $\sum_j \pi_j = 1$。
  - **物理解释：**流入状态 $j$ 的总概率速率等于流出状态 $j$ 的总概率速率（平衡方程）。
- **排队论初步 (生灭过程)：** 一种特殊的连续时间马氏链，状态只能转移到相邻状态。
  - **M/M/1 模型**是最经典的生灭过程应用，系统中的顾客数是状态。

# 6 第 6 讲：高斯过程

## 定义与核心性质

- **定义：** 一个随机过程 $\{X(t), t \in T\}$，其任意有限维分布都是**多元正态分布**。
- **核心性质：**
  1. **由二阶矩完全确定：**高斯过程的全部统计特性完全由其**均值函数** $\mu(t) = E[X(t)]$ 和**协方差函数** $C(s,t) = \mathrm{Cov}(X(s), X(t))$ 唯一确定。
  2. **平稳性关系：**宽平稳高斯过程一定是严平稳的。
  3. **线性变换封闭性：**高斯过程经过线性运算（如积分、微分）后仍然是高斯过程。

## 布朗运动 (维纳过程)

- **定义：** 布朗运动 $\{W(t), t \geq 0\}$ 是最重要的高斯过程，满足：
  1. $W(0) = 0$。
  2. 具有独立、平稳增量。
  3. 增量 $W(t) - W(s)$ 服从正态分布 $N(0, \sigma^2(t-s))$。
- **重要参数：**
  - 均值函数：$E[W(t)] = 0$。
  - 协方差函数：$\boldsymbol{C(s,t) = \sigma^2 \min(s,t)}$。
- **样本轨道性质：**处处连续，但处处不可微。

# A 附录：核心数学概念回顾

## A.1 协方差与协方差矩阵

### 协方差 (Covariance)

- **定义：** 协方差是衡量两个随机变量 $X$ 和 $Y$ **线性相关程度**的指标。其定义为：

$$\text{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])]$$

  一个更常用的计算公式是：

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$$

- **意义解释：**
  - **Cov**$(X, Y) > 0$：表示 $X$ 和 $Y$ 呈**正相关**关系。即当一个变量大于其均值时，另一个变量也倾向于大于其均值。
  - **Cov**$(X, Y) < 0$：表示 $X$ 和 $Y$ 呈**负相关**关系。即当一个变量大于其均值时，另一个变量倾向于小于其均值。
  - **Cov**$(X, Y) = 0$：表示 $X$ 和 $Y$ **不相关**。注意，不相关不等于独立，但独立一定不相关。
- **特殊情况：** 当 $Y = X$ 时，协方差就是方差。

$$\text{Cov}(X, X) = E[X^2] - (E[X])^2 = \text{Var}(X)$$

### 协方差矩阵 (Covariance Matrix)

- **定义：** 对于一个 $n$ 维随机向量 $\mathbf{X} = (X_1, X_2, \ldots, X_n)^T$，其协方差矩阵 $\boldsymbol{\Sigma}$ (或记作 $\text{Cov}(\mathbf{X})$) 是一个 $n \times n$ 的方阵，其第 $(i, j)$ 个元素为：

$$\Sigma_{ij} = \text{Cov}(X_i, X_j)$$

  矩阵形式为：

$$\boldsymbol{\Sigma} = \begin{pmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) & \ldots & \text{Cov}(X_1, X_n) \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) & \ldots & \text{Cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_n, X_1) & \text{Cov}(X_n, X_2) & \ldots & \text{Var}(X_n) \end{pmatrix}$$

- **计算公式：** 若均值向量为 $\boldsymbol{\mu} = E[\mathbf{X}]$，则

$$\boldsymbol{\Sigma} = E[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^T]$$

- **重要性质：**

  1. **对称性：** $\boldsymbol{\Sigma}$ 是一个对称矩阵，因为 $\text{Cov}(X_i, X_j) = \text{Cov}(X_j, X_i)$。

2. **半正定性：** $\mathbf{\Sigma}$ 是一个半正定矩阵。即对于任意非零常数向量 $\mathbf{c}$，有 $\mathbf{c}^T\mathbf{\Sigma}\mathbf{c} \geq 0$。这是因为 $\mathbf{c}^T\mathbf{\Sigma}\mathbf{c} = \text{Var}(\mathbf{c}^T\mathbf{X}) \geq 0$。
3. **对角线元素：**对角线上的元素是各个分量的方差。
4. **线性变换：**若对 $\mathbf{X}$ 作线性变换 $\mathbf{Y} = A\mathbf{X}$，则 $\mathbf{Y}$ 的协方差矩阵为 $\mathbf{\Sigma}_{\mathbf{Y}} = A\mathbf{\Sigma}_{\mathbf{X}}A^T$。

## A.2 矩阵的对角化

### 什么是矩阵对角化？

- **目标：** 对于一个 $n \times n$ 的方阵 $M$，找到一个可逆矩阵 $P$ 和一个对角矩阵 $D$，使得：

$$M = PDP^{-1} \quad \text{或等价地} \quad P^{-1}MP = D$$

  这个过程称为将矩阵 $M$ **对角化**。

- **意义：** 对角化可以将一个复杂的矩阵运算问题，转化为一个简单的对角矩阵运算问题。例如，计算矩阵的幂次变得非常容易：

$$M^k = (PDP^{-1})^k = (PDP^{-1})(PDP^{-1})\dots(PDP^{-1}) = PD(P^{-1}P)D\dots DP^{-1} = PD^kP^{-1}$$

  计算对角矩阵的幂次只需将对角线元素各自求幂即可。

### 如何对角化？——特征值分解

对角化的核心是找到矩阵的**特征值**和**特征向量**。

- **特征值与特征向量：** 对于矩阵 $M$，如果存在一个数 $\lambda$ 和一个非零向量 $\mathbf{v}$，使得

$$M\mathbf{v} = \lambda\mathbf{v}$$

  那么 $\lambda$ 称为 $M$ 的一个**特征值**，$\mathbf{v}$ 称为对应于 $\lambda$ 的**特征向量**。

- **对角化的条件：** 一个 $n \times n$ 矩阵 $M$ 可对角化的充要条件是它有 $n$ 个线性无关的特征向量。

- **对角化步骤：**

  1. **求特征值：**解特征方程 $\det(M - \lambda I) = 0$，求出所有特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$。
  2. **求特征向量：**对于每一个特征值 $\lambda_i$，解线性方程组 $(M - \lambda_i I)\mathbf{v} = \mathbf{0}$，求出对应的特征向量 $\mathbf{v}_i$。
  3. **构造 P 和 D：**
     - **矩阵 P：**将求出的 $n$ 个线性无关的特征向量 $\mathbf{v}_1, \dots, \mathbf{v}_n$ 按列排列构成矩阵 $P$。

$$P = [\mathbf{v}_1|\mathbf{v}_2|\dots|\mathbf{v}_n]$$

– **矩阵 D**：将与特征向量顺序对应的特征值 $\lambda_1, \dots, \lambda_n$ 放在主对角线上构成对角矩阵 $D$。

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

4. **完成对角化**：此时就有 $M = PDP^{-1}$。

### 特殊情况：对称矩阵的正交对角化

- **重要定理：** 任何一个实对称矩阵都**一定可以**被对角化。
- **正交矩阵：** 不仅如此，对称矩阵的属于不同特征值的特征向量是相互**正交**的。我们可以将这些特征向量单位化，构成一个**正交矩阵** $P$。
- **正交矩阵的性质：** $P^T P = I$，即 $P^{-1} = P^T$。
- **正交对角化：** 对于对称矩阵 $M$，我们可以找到一个正交矩阵 $P$，使得：

$$M = P\Lambda P^T \quad \text{或等价地} \quad P^T M P = \Lambda$$

其中 $\Lambda$ 是由特征值构成的对角矩阵。这在主成分分析 (PCA) 和处理协方差矩阵时非常有用。