



随机过程理论

Stochastic process theory

授课教师: 李春升教授、徐华平教授



目录

7.1 马尔可夫链的定义

7.4 遍历性与平稳分布

7.2 切普曼-柯尔莫哥洛夫方程

7.5 马尔可夫序列

7.3 马尔可夫链中的状态分类

7.6 马尔可夫过程





前言

日常生活



天气预报



球赛预测

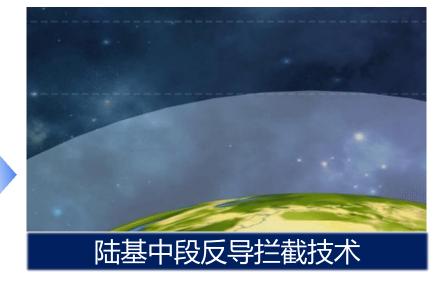


股票涨跌



语言模型

科研领域



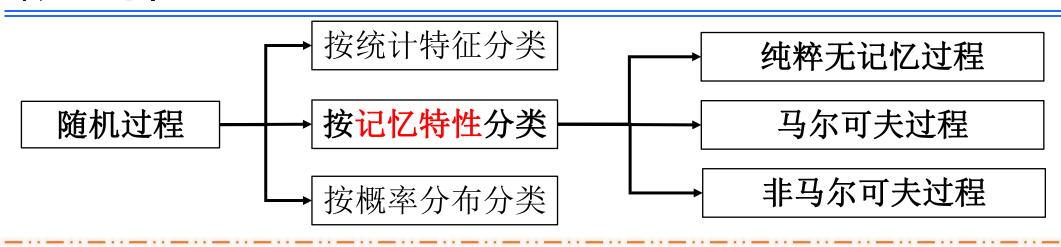
- 目标从哪来的(信息未知)
- 目标当前状态(信息可知)
- 目标下步去哪 (预测跟踪)

如何用目标当前的状态,科学合理的预测目标未来的状态?

预

测

概念引入



例: 球赛预测



过去



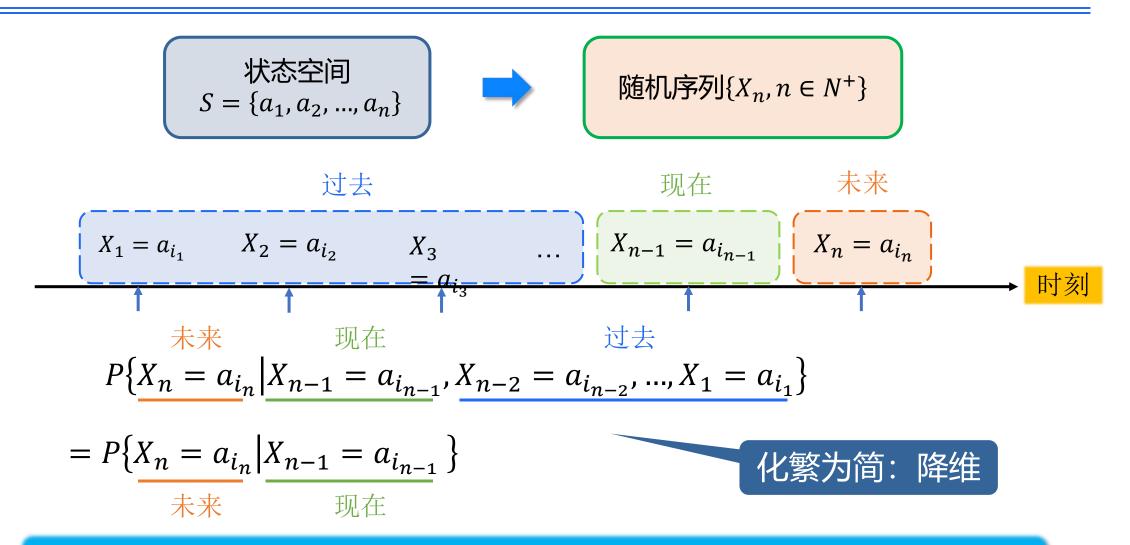
现在



未来

如何描述三者之间的关系呢?

马尔可夫过程的内涵



未来时刻状态只与当前时刻状态有关,而与过去时刻状态无关

马尔可夫链的定义

- 马尔可夫链

设 $\{X_n, n \in N^+\}$ 为一随机序列,时间参数集 $N^+ = \{0,1,2,\cdots\}$,

其状态空间 $S = \{a_1, a_2, ..., a_n\}$,若对所有的 $n \in N^+$,有

$$P\{X_n = a_{i_n} | X_{n-1} = a_{i_{n-1}}, X_{n-2} = a_{i_{n-2}}, ..., X_1 = a_{i_1}\} =$$

$$P\{X_n = a_{i_n} | X_{n-1} = a_{i_{n-1}} \}$$

则称 $\{X_n, n \in N^+\}$ 为马尔可夫链。

时间离散

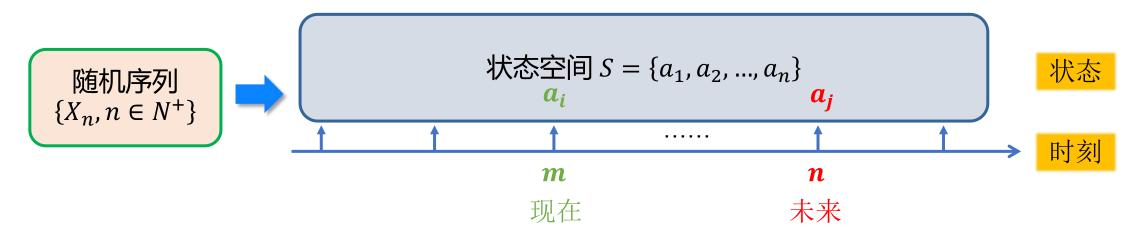
马尔可夫过程

状态离散

马尔可夫链

转移概率

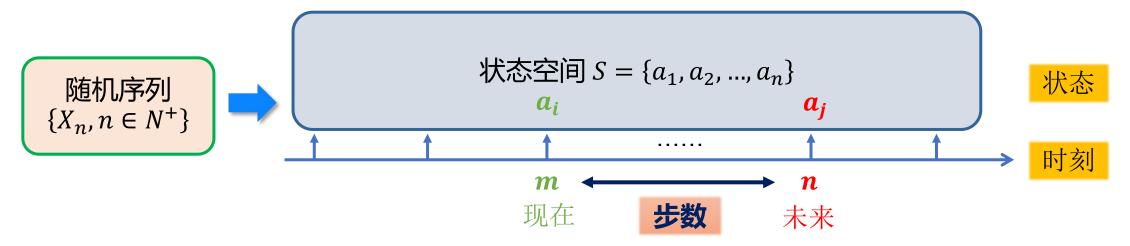
问题: 如何由当前的状态去预测未来的状态呢?



■ 转移概率

$$p_{ij}(m,n) = P\{X_n = a_j | X_m = a_i\} = P\{X_n = j | X_m = i\}$$
 时刻

转移概率



■ 基本 (一步) 转移概率

$$p_{ij}(m) = p_{ij}(m(m+1)) = P\{X_{m+1} = a_j | X_m = a_i\}$$

■ k步转移概率

$$p_{ij}^{(k)}(m) = P\{X_{m+k} = j | X_m = i\}$$

■ k步转移矩阵

$$P = \{p_{ij}^{(k)}(m), i, j \in S\}$$

例:天气预测问题

随机序列 $\{X_n, n \in N^+\}$



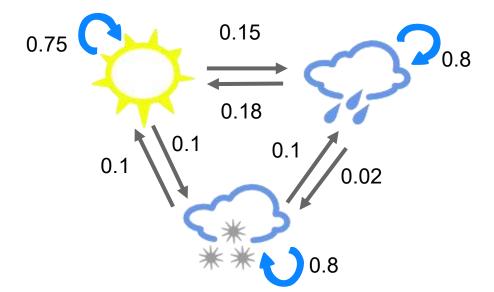
每一天的天气

状态空间 $S = \{a_1, a_2, ..., a_n\}$



晴天、下雨、下雪

问:已知今天下雪的概率为0.3,那么明天晴天的概率?



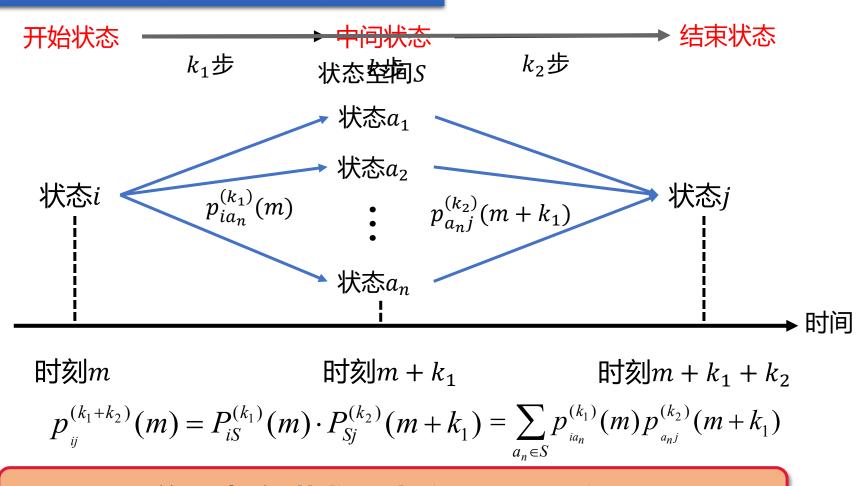
$$m{P} = egin{bmatrix} m{\pi} & m{g} & m{f} \ 0.8 & 0.02 & 0.18 \ 0.1 & 0.8 & 0.1 \ 0.15 & 0.1 & 0.75 \end{bmatrix} m{\pi} \ m{g} \ m{f} \ \m{f} \ \m$$

$$P_A = 0.3 \times 0.1 = 0.03$$



切普曼-柯尔莫戈罗夫方程

问题:如何得到k步转移概率呢?



切普曼-柯尔莫戈罗夫方程 (C-K方程)

C-K方程证明

证明
$$p_{ij}^{(m+r)}(n) = \sum_{k \in S} p_{ik}^{(m)}(n) p_{kj}^{(r)}(n+m) \quad i, j \in S$$

$$p_{ij}^{(m+r)}(n) = P\left\{X_{n+m+r} = j \middle| X_n = i\right\}$$

$$= \frac{P\left\{X_{n+m+r} = j, X_n = i\right\}}{P\left\{X_n = i\right\}}$$

$$= \sum_{k \in S} \frac{P\left\{X_{n+m+r} = j, X_{n+m} = k, X_n = i\right\}}{P\left\{X_{n+m} = k, X_n = i\right\}} \bullet \frac{P\left\{X_{n+m} = k, X_n = i\right\}}{P\left\{X_n = i\right\}}$$

$$= \sum_{k \in S} P\left\{X_{n+m+r} = j \middle| X_{n+m} = k, X_n = i\right\} \bullet P\left\{X_{n+m} = k \middle| X_n = i\right\}$$

$$= \sum_{k \in S} P\left\{X_{n+m+r} = j \middle| X_{n+m} = k\right\} \bullet P\left\{X_{n+m} = k \middle| X_n = i\right\}$$

$$= \sum_{k \in S} P\left\{X_{n+m+r} = j \middle| X_{n+m} = k\right\} \bullet P\left\{X_{n+m} = k \middle| X_n = i\right\}$$

$$= \sum_{k \in S} P\left\{i\right\} \left(n + m\right) P_{ik}^{(m)}(n) = \sum_{k \in S} P_{ik}^{(m)}(n) P_{kj}^{(r)}(n+m)$$

C-K方程证明

描述: 马尔可夫链的有限维分布由其初始分布及基本 转移概率唯一确定

$$\begin{split} &P\{X_{n_{1}}=i_{1},X_{n_{2}}=i_{2},\cdots,X_{n_{k}}=i_{k}\}\\ &=\sum_{j\in S}P\{X_{0}=j,X_{n_{1}}=i_{1},X_{n_{2}}=i_{2},\cdots,X_{n_{k-1}}=i_{k-1},,X_{n_{k}}=i_{k}\}\\ &=\sum_{j\in S}p_{i_{k-1}i_{k}}^{(n_{k}-n_{k-1})}P\{X_{0}=j,X_{n_{1}}=i_{1},X_{n_{2}}=i_{2},\cdots,X_{n_{k-1}}=i_{k-1}\}\\ &=\sum_{j\in S}p_{i_{k-1}i_{k}}^{(n_{k}-n_{k-1})}p_{i_{k-2}i_{k-1}}^{(n_{k-1}-n_{k-2})}P\{X_{0}=j,X_{n_{1}}=i_{1},X_{n_{2}}=i_{2},\cdots,X_{n_{k-2}}=i_{k-2}\}=\cdots\cdots\\ &=\sum_{j\in S}p_{i_{k-1}i_{k}}^{(n_{k}-n_{k-1})}p_{i_{k-2}i_{k-1}}^{(n_{k-1}-n_{k-2})}\cdots p_{i_{1}i_{2}}^{(n_{2}-n_{1})}P\{X_{n_{1}}=i_{1},X_{0}=j\}\\ &=\sum_{j\in S}p_{i_{k-1}i_{k}}^{(n_{k}-n_{k-1})}p_{i_{k-2}i_{k-1}}^{(n_{k-1}-n_{k-2})}\cdots p_{i_{1}i_{2}}^{(n_{2}-n_{1})}p_{ji_{1}}^{(n_{1})}P\{X_{0}=j\} \end{split}$$

> 举例: 随机游走

齐次马尔可夫链

• 齐次马尔可夫链

如果在马尔可夫链中

$$p_{ij}(m) = P\{X_{m+1} = a_j | X_m = a_i\} = p_{ij}$$
与时刻m无关

即从状态*i*转移到状态*j*的概率与时刻*m*无关,则称这类马尔可夫为齐次马尔可夫链,它具有平稳转移概率的特性

齐次马尔可夫链的C-K方程

齐次形式
$$p_{ij}^{(k_1+k_2)}(m) = p_{ij}^{(k_1+k_2)} = \sum_{k \in S} p_{ik}^{(k_1)} p_{kj}^{(k_2)}$$

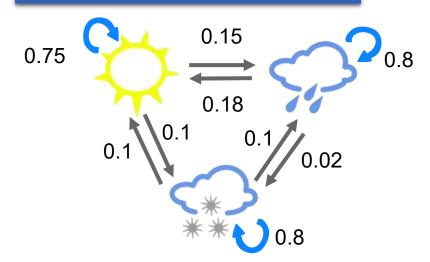
矩阵形式

$$P^{(k)} = P^{(k_1+k_2)} = P^{(k_1)}P^{(k_2)}$$

= $P^{(k_1)}P^{(k_2-1)}P = \cdots = P^k$

可以完全确定

例:天气预测问题续



问:已知今天下雪的概率是0.3,

那么今天下雪五天后是晴天的概率?

$$\mathbf{P}^{(5)} = \mathbf{P}^5 = \begin{bmatrix} 0.4854 & 0.1354 & 0.3793 \\ 0.3045 & 0.401 & 0.2945 \\ 0.3574 & 0.2201 & 0.4225 \end{bmatrix}$$
晴 $P_A = 0.3 \times 0.2945 = 0.08835$

目标跟踪-目标状态建模

问题:如何进行目标预测跟踪?

假设目标做一维匀速直线运动,

目标初始位置为 x_0 ,目标速度为v

$$x = x_0 + v \cdot t$$
 牛顿运动方程



假设对目标运动情况进行离散采样(时间间隔为T),则

$$x_1 = x_0 + v \cdot T$$

$$x_2 = x_1 + v \cdot T \qquad \qquad x_{k+1} = x_k + v \cdot T$$

马尔可夫链

.

第k + 1时刻目标的位置 x_{k+1} 由k时刻目标位置 x_k 、速度v以及采样间隔T决定

目标跟踪-目标状态建模

牛顿运动方程



离散状态模型

$$x = x_0 + v \cdot t$$

$$x_{k+1} = x_k + v_k \cdot T$$

目标运动状态(一维)可构建为 $X_k = [x_k, v_k]^T$, 通常称为**系统状态**

离散状态模型可建模为 $X_{k+1} = \mathbf{F} \cdot X_k$,实际中考虑运动过程噪声可建模为

$$X_{k+1} = \mathbf{F}X_k + W_k$$

其中
$$\mathbf{F} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}$$
, $W_k = \begin{bmatrix} \mathbf{w}_{k,x} \\ \mathbf{w}_{k,v} \end{bmatrix}$ 为服从高斯分布的过程噪声

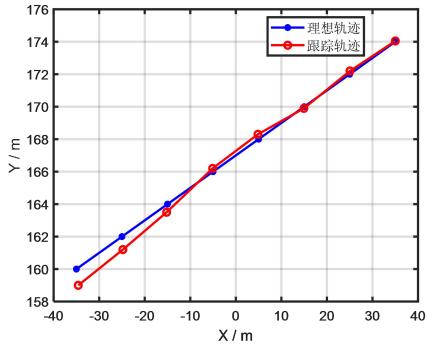
目标跟踪-目标状态建模

$$X_{k+1} = \mathbf{F}X_k + W_k$$

通常目标在三维空间中运动,目标运动模型为
$$X_{k+1} = \mathbf{F} X_k + W_k \qquad \mathbf{F} = \begin{bmatrix} 1 & T & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & T & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & T \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$
目标运动状态可构建为 $\mathbf{X}_k = [x_k, v_k, v_k, v_k, v_k, v_k, v_k, v_k]^T$

目标运动状态可构建为 $X_k = [x_k, v_x, y_k, v_y, z_k, v_z]^T$

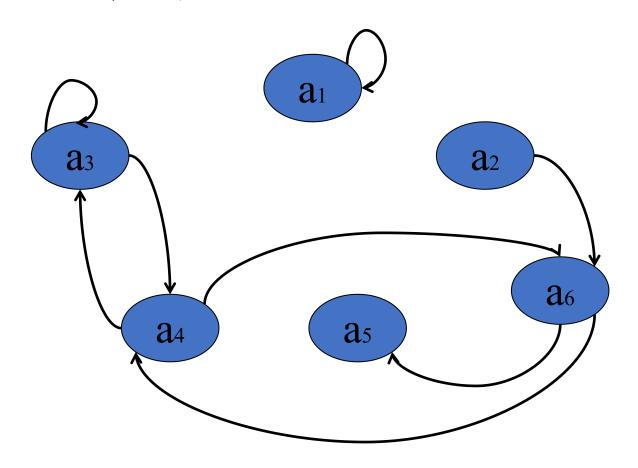






- ✓ 状态可达 由状态i到状态j, $P_{ij}^{(k)} > 0$, 记为 $i \rightarrow j$
- \checkmark 定理 由 $i\rightarrow k$, $k\rightarrow j$, 则 $i\rightarrow j$
- ✓ 状态相通 状态 $i \rightarrow$ 状态j,且状态 $j \rightarrow$ 状态I,记为 $i \leftrightarrow j$
- \checkmark 定理 由状态 $i \leftrightarrow$ 状态k,状态 $k \leftrightarrow$ 状态j,则状态 $i \leftrightarrow$ 状态j

例: 试分析下图:



2、首次进入时间

对于任何两个状态i和j,在事件 $\{X_0=i\}$ 上引入随机变量T

$$T_{ij}(\omega) \triangleq \min\{n : X_0(\omega) = i, X_n(\omega) = j, n \ge 1\}$$

$$T_{ij}$$
 从状态 i 出发,首次进入状态 j 的时间 使 $X_n = j$ 的最小正值 n

$$T_{ij}(\omega) = \infty$$
 终身等待

2、首次进入时间

定义:从状态i经过n步首次进入状态j的概率

 $0 \le f_{ii}^{(n)} \le f_{ij} \le 1$

$$f_{ij}^{(n)} = P\{T_{ij} = n \mid X_0 = i\}$$

对于 $n \in N_{\infty}$
有 $f_{ij} = \sum_{n < \infty} f_{ij}^{(n)} = \sum_{n < \infty} P\{T_{ij} = n \mid X_0 = i\}$
 $= P\{T_{ij} < \infty\}$
 $f_{ij}^{(\infty)} = P\{T_{ij} = \infty\} = 1 - f_{ij}$

2、首次进入时间

定理1:
$$p_{ij}^{(n)} = \sum_{v=1}^{n} f_{ij}^{(v)} p_{jj}^{(n-v)}$$
证明
$$p_{ij}^{(n)} = P\{X_n = j \mid X_0 = i\}$$

$$= P\{T_{ij} \le n, X_n = j \mid X_0 = i\}$$

$$= \sum_{v=1}^{n} P\{T_{ij} = v, X_n = j \mid X_0 = i\}$$

$$= \sum_{v=1}^{n} P\{T_{ij} = v \mid X_0 = i\} P\{X_n = j \mid X_0 = i, T_{ij} = v\}$$

$$= \sum_{v=1}^{n} P\{T_{ij} = v \mid X_0 = i\} P\{X_n = j \mid X_0 = i, X_1 \neq j, \cdots X_{v-1} \neq j, X_v = j\}$$

$$= \sum_{v=1}^{n} P\{T_{ij} = v \mid X_0 = i\} P\{X_n = j \mid X_v = j\} = \sum_{v=1}^{n} f_{ij}^{(v)} p_{jj}^{n-v}$$

2、首次进入时间

定理2:
$$f_{ij} > 0$$
 \longleftrightarrow $i \to j$
推论: $i \longleftrightarrow j$ \longleftrightarrow $f_{ij} > 0$ 且 $f_{ji} > 0$
当 $i = j$ 时
 $f_{ii} = 1$ \longleftrightarrow 状态i常返
 $f_{ij} < 1$ \longleftrightarrow 状态i非常返,滑过的

2、首次进入时间

定理3:

如果状态j是常返的,则以概率1系统无穷次返回状态j; 如果状态j是非常返的,则以概率1系统只有有限次返回 状态j。亦即系统无穷次返回状态j的概率为零。

推论:

$$Q_{ij} = \begin{cases} f_{ij} & j$$
为常返态
$$0 & j$$
为非常返态

2、首次进入时间

定理4:

状态
$$j$$
为常返的充要条件是 $\sum_{n=1}^{\infty} p_{jj}^{(n)} = \infty$

如果状态j为非常返的,则 $\sum_{n=0}^{\infty} p_{jj}^{(n)} = \frac{1}{1-f_{ii}} < \infty$

平均返回时间 $m_j = \sum_{j=1}^{\infty} n f_{jj}^{(n)}$

$$m_j < \infty$$
 正常返态

$$m_j < \infty$$
 正常返态
$$m_j = \infty / \frac{1}{m_j} = 0$$
 零常返态

2、首次进入时间

定理5:

如果状态*i*, *j*是马尔可夫链两个相通的状态,则它们同为常返态,或同为非常返态。

定理6:

如果状态j是非常返的,则对每一个状态i, $\sum\limits_{n=0}^{\infty}p_{ij}^{(n)}<\infty$ 且对每一个状态i, $\lim\limits_{n\to\infty}p_{ij}^{(n)}=0$

3、状态空间的分解

定义:

由一些状态组成的集合C,如果对任意的 $i \in C$, $j \notin C$ 自状态i出发,不能到达状态j,则称状态集合C为闭集。

- 1、若单个状态形成一个闭集,则称这个闭集为吸收状态。
- 2、在一个闭集内,若不包含任何子闭集,则称该闭集为 不可约的,这时所有状态之间都是相通的。

3、状态空间的分解

定理1:

在转移概率矩阵P(n)中,仅保留同类中各状态间的转移概率,将其他所有行和列都删去,则剩下一个随机矩阵,其中基本关系仍满足

$$\sum_{k \in C} p_{ik} = 1 \qquad i \in C \qquad p_{ij}^{(n+m)} = \sum_{r \in C} p_{ir}^{(n)} p_{rj}^{(m)}$$

这意味着有一定义在C上的马尔可夫子链,且这个子链可以不涉及所有其他状态而被独立地研究。

3、状态空间的分解

定理2:

所有常返状态构成一个闭集C。

定理3:

在一个马列尔可夫链中,所有常返状态可以分为若干个互不相交的闭集 $\{C_n,n=1,2,3,\cdots\}$,且有:

- (1) C_k中任二状态相通
- (2) C_k 中的任一状态和 C_m 中的任一状态,在 $k \neq m$ 时互不相通。

3、周期状态和非周期状态

定义:

称正整数集合 $\{n: n \geq 1, p_{ii}^{(n)} > 0\}$ 的最大公约数 $G.C.D\{n: n \geq 1, p_{ii}^{(n)} > 0\}$ 为状态i的周期,记之为d。或者说,状态i是具 周期d的周期性状态。

特别地, 当d=1时, 则称状态i是无周期的。

当 $\{n: n \ge 1, p_{ii}^{(n)} > 0\}$ 为空集时,不考虑i的周期。

非周期的正常返态称为遍历状态。



遍历性与平稳分布

1、遍历性

 \checkmark 定义 设 $\{X_n=i\}$ 为齐次马氏链,对一切状态i,j有

$$\lim_{n\to\infty}p_{ij}^{(n)}=\pi_j$$

其中,πi为平稳分布

$$\rightarrow$$
 $\pi_j \geqslant 0$

$$\pi_{j} \ge 0$$

$$\sum_{j} \pi_{j} = 1$$

遍历性与平稳分布

 \checkmark 定理 对有限状态的齐次马氏链,若存在一个正整数 \mathbf{m} ,使得对一切状态i,j有 $p_{ij}^{(m)} > 0$,则

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j$$

且该马氏链是遍历的

$$\rightarrow$$
 $\pi_j \geqslant 0$

$$\sum_{j} \pi_{j} = 1$$

遍历性与平稳分布

2、平稳分布

 \checkmark 定义 设 $\{X_n=i\}$ 为齐次马氏链,概率分布 p_i ,满足

$$p_{j} = \sum_{i} p_{i} p_{ij} \quad j = 0, 1, 2, \cdots$$

其中, p_{j} 为平稳分布

- $\rightarrow p_j \geqslant 0$
- $\sum_{j} p_{j} = 1$
- > $n \to \infty$ by, $p_j = \pi_j$

遍历性与平稳分布

对于平稳分布,有

$$p_{j} = \sum_{i} p_{i} p_{ij} = \sum_{i} (\sum_{k} p_{k} p_{ki}) p_{ij} = \sum_{k} p_{k} (\sum_{i} p_{ki} p_{ij})$$
$$= \sum_{k} p_{k} p_{kj}^{(2)} = \dots = \sum_{k} p_{k} p_{kj}^{(n)}$$

$$ightharpoonup$$
 另外 $\pi = [\pi_1 \ \pi_2 \ \cdots \ \pi_n]$ $P = [p_{ij}]_{N \times N}$ 则 $\pi = \pi P$



马尔可夫序列

\checkmark 定义 设一个随机序列 X_n 连续,有

$$F(x_n | x_{n-1}, \dots, x_1) = F(x_n | x_{n-1})$$

则称此随机序列为马尔可夫序列

其联合概率密度为

$$f(x_n | x_{n-1}, \dots, x_1) = f(x_n | x_{n-1})$$

- \triangleright 齐次性 $f(x_n | x_{n-1})$ 与n无关
- \rightarrow 平稳性 对齐次过程有 $f(x_n) = f(x)$

马尔可夫序列

> 切普曼—柯尔莫哥洛夫方程

$$f(x_n | x_s) = \int f(x_n | x_r) f(x_r | x_s) dx_r$$

> 马尔可夫序列中,若已知现在,则过去和将来独立

$$f(x_n, x_s | x_r) = f(x_n | x_r) f(x_s | x_r)$$
 $x_s < x_r < x_n$

> 马尔可夫序列的逆也具有马尔可夫特性

$$f(x_1 | x_2, \dots, x_n) = f(x_1 | x_2)$$

$$f(x_n, x_s \mid x_r) = f(x_n \mid x_r) f(x_s \mid x_r)$$

证:

$$f(x_n, x_s | x_r) = \frac{f(x_n, x_s, x_r)}{f(x_r)}$$

$$= \frac{f(x_n | x_r) f(x_r | x_s) f(x_s)}{f(x_r)}$$

$$= \frac{f(x_n | x_r) f(x_r, x_s)}{f(x_r)}$$

$$= f(x_n | x_r) f(x_s | x_r)$$

$$f\left(x_{n} \mid x_{n+1}, x_{n+2}, \dots, x_{n+k}\right) = f\left(x_{n} \mid x_{n+1}\right)$$

$$f\left(x_{n} \mid x_{n+1}, x_{n+2}, \dots, x_{n+k}\right) = \frac{f\left(x_{n}, x_{n+1}, \dots, x_{n+k}\right)}{f\left(x_{n+1}, \dots, x_{n+k}\right)}$$

$$= \frac{f\left(x_{n}\right) f\left(x_{n+1} \mid x_{n}\right) f\left(x_{n+2} \mid x_{n+1}\right) \dots f\left(x_{n+k} \mid x_{n+k-1}\right)}{f\left(x_{n+1}\right) f\left(x_{n+2} \mid x_{n+1}\right) \dots f\left(x_{n+k} \mid x_{n+k-1}\right)}$$

$$= \frac{f\left(x_{n}\right) f\left(x_{n+1} \mid x_{n}\right)}{f\left(x_{n+1}\right)} = \frac{f\left(x_{n+1}, x_{n}\right)}{f\left(x_{n+1}\right)}$$

$$= f\left(x_{n} \mid x_{n+1}\right)$$



- > 定义
 - ✓ 随机过程 $\{X(t), t \in T\}$
 - $\checkmark F\{x_n, t_n | x_{n-1}, t_{n-1}; x_{n-2}, t_{n-2}; \dots; x_1, t_1\} = F\{x_n, t_n | x_{n-1}, t_{n-1}\}$
- > 理解
 - ✓ 有两个参量
 - √ 概率密度函数若存在,则

$$f(x_n, t_n | x_{n-1}, t_{n-1}; x_{n-2}, t_{n-2}; \dots; x_1, t_1) = f(x_n, t_n | x_{n-1}, t_{n-1})$$

- ✓ 转移概率分布和转移概率密度函数
- ✓ 齐次马尔可夫过程

$$F\{y,t|x,s\} = F\{y,t-s|x\}$$

> 切普曼—柯尔莫哥洛夫方程

$$f(x_n, t_n | x_s, t_s) = \int f(x_n, t_n | x_r, t_r) f(x_r, t_r | x_s, t_s) dx_r$$

马氏过程的联合概率密度完全由一维和二维分布函数描述

$$f(x_n, t_n; \dots; x_1, t_1) = \frac{\sum_{k=2}^{n} f(x_k, t_k; x_{k-1}, t_{k-1})}{\sum_{k=2}^{n-1} f(x_k, t_k)}$$

> 马氏过程的逆也具有马尔可夫特性

$$f(x_1,t_1|x_2,t_2;\dots;x_n,t_n) = f(x_1,t_1|x_2,t_2)$$

证:

$$f(x_{n}, t_{n}; \dots; x_{1}, t_{1}) = \frac{\sum_{k=2}^{n} f(x_{k}, t_{k}; x_{k-1}, t_{k-1})}{\sum_{k=2}^{n-1} f(x_{k}, t_{k})}$$

$$f(x_{1}, t_{1}; \dots; x_{n}, t_{n}) = f(x_{n}, t_{n} \mid x_{n-1}, t_{n-1}; \dots; x_{1}, t_{1})$$

$$f(x_{n-1}, t_{n-1} \mid x_{n-2}, t_{n-2}; \dots; x_{1}, t_{1}) \dots f(x_{1}, t_{1})$$

$$= f(x_{1}, t_{1}) \prod_{k=1}^{n-1} f(x_{k+1}, t_{k+1} \mid x_{k}, t_{k})$$

$$= \frac{\prod_{k=1}^{n-1} f_{2}(x_{k+1}, t_{k+1}; x_{k}, t_{k})}{\prod_{k=2}^{n-1} f(x_{k}, t_{k})}$$

作业

结合在工作或生活中的实例,给出随机过程在实际中的具体应用

作业

7.1, 7.2, 7.4, 7.7, 7.9

悟已往之不谏, 知来者之可追。



---陶渊明